

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 ; Search time 43.5294 Seconds  
(without alignments)  
72.928 Million cell updates/sec

Title: US-09-171-432a-39  
Perfect score: 107  
Sequence: 1 SHICRRPKYKELRLEVGKQR 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

al number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneq-emb1/AA2000.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneq-emb1/AA2001.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneq-emb1/AA2002.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneq-emb1/AA2003.DAT.\*  
24: /SIDSL/gcgdata/geneseq/geneq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	107	100.0	20	18 AAM42922
2	107	100.0	21	22 AAB6439
3	107	100.0	366	6 AAP50230
4	107	100.0	854	6 AAP50287
5	107	100.0	993	6 AAP50116
6	107	100.0	993	6 AAP50231
7	107	100.0	1077	20 AAW95559
8	107	100.0	1091	14 AAR33426
9	107	100.0	2227	11 AAR05697

10	107	100.0	2227	18 AAW34074
11	107	100.0	2227	21 AAB1607
12	107	100.0	2227	21 AAB1608
13	107	100.0	2227	21 AAB1609
14	107	100.0	2227	23 ABB31727
15	107	100.0	2227	23 ABB31728
16	107	100.0	2227	23 ABB31729
17	107	100.0	2227	23 ABE19899
18	107	100.0	2227	24 ABU08639
19	107	100.0	2227	24 ABU08640
20	107	100.0	2227	24 ABU08641
21	91	85.0	2227	7 ABE0066
22	84.5	79.0	839	12 AAR15629
23	67	62.6	20	18 AAB42923
24	67	62.6	21	22 AAB69440
25	45.5	42.5	178	22 AAU03662
26	45.5	42.5	330	23 ABB30218
27	45.5	42.5	341	23 ABB25499
28	45	42.1	71	23 ABB62888
29	44	41.1	71	23 AAG07411
30	44	41.1	88	21 AAG61775
31	44	41.1	211	22 AAG81246
32	44	41.1	346	12 AAR11068
33	43	40.2	45	22 AAB84231
34	43	40.2	45	22 AAE13255
35	43	40.2	52	22 AAB84230
36	43	40.2	52	22 AAE13254
37	43	40.2	287	21 AAG06837
38	43	40.2	312	21 AAG06837
39	43	40.2	320	21 AAG06836
40	43	40.2	332	23 ABB91335
41	43	40.2	372	22 ABB71552
42	43	40.2	391	8 AAE70479
43	43	40.2	391	19 AAB25305
44	43	40.2	391	19 AAW50001
45	43	40.2	413	22 AAB84220

## ALIGNMENTS

RESULT 1  
ID AAM42922 standard; peptide; 20 AA.  
AC AAM42922;  
XX  
DE 28-APR-1998 (first entry)  
XX  
XX Immunogenic Hepatitis A virus peptide YK-1315.  
DE  
XX Immunogenic peptide; immunogenic epitope; P2A protein;  
KW Immune response; antibody.  
XX  
OS Synthetic.  
OS Hepatitis A virus.  
XX  
PN WO9740147-A1.  
PD 30-OCT-1997.  
XX  
PF 18-APR-1997; 97WO-US06891.  
XX  
PR 19-APR-1996; 96US-0015644.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PI Fields HA, Khudyakov YE.  
XX  
DR WPI: 1997-535831/49.  
XX  
PT Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an  
PT Immune response to HAV in a mammal or to detect the presence of

PT antibodies against HAV in a mammal  
 XX  
 PS Claim 18; Page 112; 140pp; English.  
 XX  
 CC Peptides AAW42922-30 are immunogenic peptides corresponding to  
 CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are  
 CC substantially similar to a portion of the amino acid sequence of the P2A  
 CC protein of HAV corresponding to amino acids 792-980. The present peptide  
 CC is derived from amino acids 792-811, and has a reactivity of 54.2% with  
 CC acute sera. Compositions containing the peptides can be used to induce an  
 CC immune response to HAV in a mammal. The peptides can also be used to  
 CC detect the presence of antibodies against HAV in mammalian serum. The  
 CC peptides can also be used to make an antibody against HAV by  
 CC administering the peptide to a mammal.  
 CC  
 SQ Sequence 20 AA;  
 Query Match 100.0%; Score 107; DB 18; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-10;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 SHIECRKPYKELRLEVGKOR 20  
 1 SHIECRKPYKELRLEVGKOR 20  
 Db  
 RESULT 2  
 AAB69439  
 ID AAB69439 standard; Peptide; 21 AA.  
 XX  
 AC AAB69439;  
 XX  
 DF 20-APR-2001 (first entry)  
 XX  
 DE Synthetic HAV P2A peptide, SEQ ID NO: 39.  
 XX  
 KW Hepatitis A virus; HAV, immunogen; immunostimulant; virucide; vaccine;  
 KW antigen; major structural capsid polypeptide; HAV antibody detection.  
 XX  
 OS Hepatitis A virus.  
 OS Synthetic.  
 XX  
 PN WO200105824-A2.  
 XX  
 PD 25-JAN-2001.  
 XX  
 PF 14-JUL-2000; 2000WO-US19267.  
 XX  
 PY 15-JUL-1999; 99US-0144412.  
 XX  
 (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Fields HA, Khudyakov YE;  
 DR WPI; 2001-112681/12.  
 XX  
 PT Synthetic peptides used as antigen sources for enzyme immunoassays  
 PT detecting anti-hepatitis A virus and as vaccines -  
 XX  
 PS Claim 13; Page 93; 130pp; English.  
 XX  
 CC The present sequence is one of a number of synthetic peptides which are  
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides  
 CC comprise antigenic epitopes of the major structural capsid polypeptides  
 CC or non-structural polypeptides of HAV with one or more glutamine  
 CC molecules at the carboxy end of the peptide. The peptides are used to  
 CC detect the presence of antibodies against HAV in mammalian serum, to  
 CC detect the presence of HAV in a human or animal through the binding of  
 CC the peptide to an antibody, to detect acute phase infection by detecting  
 CC IgM antibodies in mammalian serum and detecting convalescence in a  
 CC mammal. The peptides are used to detect or quantify HAV antibodies in  
 CC samples in clinical or research-based assays using immunoblotting,  
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,

CC tracking of radioactive or bioluminescent markers, chromatography or  
 CC electrophoresis. The peptides are used to induce an immune response to  
 CC HAV when administered to a human or animal. Glutamine at the carboxy  
 CC end of the peptides enhances the IgM antibody reactivity.  
 CC  
 SQ Sequence 21 AA;  
 Query Match 100.0%; Score 107; DB 22; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 1e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 SHIECRKPYKELRLEVGKOR 20  
 1 SHIECRKPYKELRLEVGKOR 20  
 Db  
 RESULT 3  
 AAP50230  
 ID AAP50230 standard; Protein; 366 AA.  
 XX  
 AC AAP50230;  
 XX  
 DF 28-NOV-1991 (first entry)  
 XX  
 DE Sequence of hepatitis A virus (HAV) surface protein (VP-1).  
 XX  
 KW Hepatitis A virus vaccine; immunisation; monoclonal antibody;  
 KW diagnostic assay.  
 XX  
 OS Hepatitis A virus.  
 XX  
 PN EPI38704-A.  
 XX  
 PD 24-APR-1985.  
 XX  
 PF 09-OCT-1984; 84EP-0402025.  
 XX  
 PR 02-MAR-1984; 84US-0585942.  
 PR 14-OCT-1983; 83US-0541836.  
 XX  
 PA (MERI ) MERCK & CO INC.  
 XX  
 PI Hughes JV, Scolnick EM, Tomassini JE;  
 XX  
 DR WPI; 1985-100818/17.  
 DR N-PSDB; AAN50274.  
 XX  
 PT New hepatitis A virus surface protein - useful for binding to  
 PT neutralising antibodies to the virus  
 XX  
 PS Claim 21; Page 46-48; 49pp; English.  
 XX  
 CC VP1 is isolated by solubilisation of the intact virus in an aq.  
 CC anionic surfactant and a reducing agent. The viral proteins are sepd.  
 CC and the protein of molecular wt. 33000 daltons is sepd.  
 XX  
 SQ Sequence 366 AA;  
 Query Match 100.0%; Score 107; DB 6; Length 366;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 SHIECRKPYKELRLEVGKOR 20  
 1 SHIECRKPYKELRLEVGKOR 320  
 Db  
 RESULT 4  
 AAP50287  
 ID AAP50287 standard; Protein; 854 AA.  
 XX  
 AC AAP50287;  
 XX

```

DT 25-MAR-2003 (updated)
DT 30-NOV-1991 (first entry)
XX
XX
DE Sequence encoded by hepatitis A virus (HAV) cDNA from near the
DE genome 5' terminus to the end of the area corresponding to the
DE capsid protein region of poliovirus RNA.
XX
XX Hepatitis A virus assay; antigen; antibody.
XX
XX Hepatitis A virus.
XX
XX W06501517-A.
XX
XX 11-APR-1985.
XX
XX 27-SEP-1984; 84MO-US01552.
XX
XX 30-SEP-1983; 83US-0537911.
XX
XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;
XX Racanello VR.
XX
XX WPI: 1985-098846/16.
XX N-PSDB; AAN50330.
XX
XX New hepatitis A virus cDNA - useful in assays for the virus and
XX for prodn. of the viral antigen and antibodies to it
XX
XX Example: Fig 7; 60pp; English.
XX
XX The inventors claim HAV cDNA and a method for producing it, whereby
XX large amts. can be obtd. economically. The cDNA is useful in the
XX assay for detection of HAV quickly and easily and with high
XX sensitivity and specificity. The HAV cDNA is also used in the prodn.
XX of HAV antigen or antibodies to it. The antibodies may be monoclonal.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 854 AA;
SQ
Query Match 100.0%; Score 107; DB 6; Length 854;
Best Local Similarity 100.0%; Pred. No. 5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPYKELREVGKOR 20
DB 792 SHIECRKPYKELREVGKOR 811
RESULTS 5
AAP50116
ID AAP50116 standard; Protein: 993 AA.
XX
XX AAP50116;
XX
XX 25-MAR-2003 (updated)
XX 30-SEP-1991 (first entry)
XX
XX Sequence of Hepatitis A virus (HAV) immunogenic peptides
XX VP-1, VP-2, VP-3 and VP-4.
XX
XX Antigenic protein; immunogen; vaccine.
XX
XX Hepatitis A virus (strain CR326).
XX
XX EPI54587-A.
XX
XX 11-SEP-1985.
XX
XX 27-FEB-1985; 85EP-0400369.
XX
XX 02-MAR-1984; 84US-0585818.
XX

```

```

XX
XX (MERI ) MERCK & CO INC.
XX
XX Linemeyer DL, Menke JG, Rueben RG, Mltra SW;
XX
XX WPI: 1985-224964/37.
XX N-PSDB; AAN50139.
XX
XX New nucleotide sequences coding for hepatitis A virus antigens -
XX useful for eliciting normal immune response and in vaccines for
XX protecting against the virus
XX
XX Example: Page 11-17; 32pp; English.
XX
XX Within the sequence in AAN50139 is encoded the information necessary
XX to make the antigenic proteins of HAV. The sequences encoding for
XX the structural proteins begin at base 403. The key sub-unit
XX sequences within VP-1, designated Sequences I, II, III, IV, and V,
XX start, respectively at 1882, 1963, 1999, 2146, 2347. Other
XX nucleotide sequences which are valuable as encoding antigenic
XX proteins are the sequences from base 1749 to base 2722; from base
XX 1487 to base 2980 and from base 1644 to base 2722. The sequence from
XX base 1749 to base 2722 is esp. valuable as a vector for producing
XX antigen protein. Sequences II-V are claimed. X in AAP50116 denotes the
XX translation of a stop codon.
XX (Updated on 25-MAR-2003 to correct PI field.)
XX
XX Sequence 993 AA;
SQ
Query Match 100.0%; Score 107; DB 6; Length 993;
Best Local Similarity 100.0%; Pred. No. 5.9e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPYKELREVGKOR 20
DB 928 SHIECRKPYKELREVGKOR 947
RESULTS 6
AAP50231
ID AAP50231 standard; Protein: 993 AA.
XX
XX AAP50231;
XX
XX 28-NOV-1991 (first entry)
XX
XX Sequence encoded by partial sequence of hepatitis A virus (HAV),
XX including surface protein (VP-1).
XX
XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;
XX diagnostic assay.
XX
XX Hepatitis A virus.
XX
XX Key Location/Qualifiers
XX FH 628..993
XX FT /note="claimed; X denotes translated stop codons
XX FT and unspecified triplets"
XX
XX EPI38704-A.
XX
XX 24-APR-1985.
XX
XX 09-OCT-1984; 84EP-0402025.
XX
XX 02-MAR-1984; 84US-0585942.
XX 14-OCT-1983; 83US-0541836.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Hughes JV, Scolnick EM, Tomassini JE;
XX
XX WPI: 1985-100818/17.
XX

```

DR N-PSDB; AAN50274.  
 XX  
 PT New hepatitis A virus surface protein - useful for binding to  
 PT neutralising antibodies to the virus  
 XX  
 PS Disclosure; Page 17-23; 49pp; English.  
 CC  
 CC WPI is isolated by solubilisation of the intact virus in an aq.  
 CC anionic surfactant and a reducing agent. The viral proteins are sepd.  
 CC and the protein of molecular wt. 33000 daltons is sepd.  
 XX  
 SQ Sequence 993 AA;  
 Query Match 100.0%; Score 107; DB 6; Length 993;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SHIECRKPKELRLVGVKOR 20  
 |||||  
 928 SHIECRKPKELRLVGVKOR 947  
 RESULT 7  
 AAM95559 ID AAM95559 standard; Protein; 1077 AA.  
 AC AAM95559;  
 XX  
 DT 28-APR-1999 (first entry)  
 XX  
 DE A partial hepatitis A virus (HAV) protein.  
 XX  
 KW Hepatitis A virus protein; HAV; P2 region;  
 KW cell-culture-adapted HAV strain; infection; accelerated growth.  
 XX  
 OS Hepatitis A virus.  
 XX  
 PN US5849562-A.  
 XX  
 PD 15-DEC-1998.  
 XX  
 PF 06-JUN-1995; 95US-0468926.  
 XX  
 PR 06-NOV-1991; 91US-0788262.  
 PR 30-SEP-1983; 83US-0537911.  
 PR 27-SEP-1984; 84US-0654942.  
 PR 06-OCT-1988; 88US-0256135.  
 PR 06-JUN-1995; 95US-0468926.  
 (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PI Emerson SU, Purcell RH;  
 DR WPI; 1999-094412/08.  
 DR N-PSDB; AAX01006.  
 XX  
 PT Chimeric hepatitis A virus strains - with P2 region from  
 PT cell-culture-adapted strain in wild-type genome  
 XX  
 PS Disclosure; Fig 7A-L; 36pp; English.  
 CC  
 CC The present sequence represents a partial hepatitis A virus (HAV)  
 CC protein. The specification describes a DNA construct consisting  
 CC of a wild-type HAV genome in which the P2 region is replaced by the  
 CC P2 region from a cell-culture-adapted HAV strain. The construct is  
 CC used to demonstrate that mutations in the P2 region of a  
 CC cell-culture-adapted HAV strain are sufficient for establishment of  
 CC infection and accelerated growth in cell culture.  
 CC  
 SQ Sequence 1077 AA;  
 Query Match 100.0%; Score 107; DB 20; Length 1077;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SHIECRKPKELRLVGVKOR 20  
 |||||  
 1015 SHIECRKPKELRLVGVKOR 1034  
 Db  
 RESULT 8  
 AAR32426 ID AAR32426 standard; Protein; 1091 AA.  
 XX  
 AC AAR32426;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 17-DEC-2001 (updated)  
 DT 10-JUN-1993 (first entry)  
 XX  
 DE Translated from 5' region of Hepatitis A virus genomic clone.  
 XX  
 KW HAV HM-175; chronic liver disease; picornavirus.  
 XX  
 OS Hepatitis A virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 238..1091  
 FT /label= ORF  
 FT /note= "second putative initiation codon at  
 FT position 240"  
 FT Region 1..711  
 FT /note= "X's correspond to nonsense codons,  
 FT 1.e. this region is not an ORF"  
 XX  
 PN USN7788262-N.  
 XX  
 PD 15-DEC-1992.  
 XX  
 PF 30-SEP-1983; 83US-0536911.  
 XX  
 PR 27-SEP-1984; 84US-0654942.  
 PR 06-OCT-1988; 88US-0256135.  
 PR 30-SEP-1983; 83US-0536911.  
 PR 06-NOV-1991; 91US-0788262.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
 PI Ticehurst JR, Baltimore D, Felstone SM, Purcell RH, Racanietello VR;  
 PI Baroudy BM, Emerson SU;  
 XX  
 DR WPI; 1993-067429/08.  
 DR N-PSDB; AAQ36934.  
 XX  
 PT Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn.  
 PT of antigen and antibodies  
 XX  
 PS Disclosure; Fig 7; 65pp; English.  
 CC  
 CC HAV virion RNA was extracted from the livers of marmosets which had  
 CC been inoculated with HAV (the HAV had previously been passaged twice  
 CC in marmosets). The RNA was used to prepare ds cDNA clones by  
 CC standard methods. Clones contg. inserts which hybridised to RNA from  
 CC HAV-infected African Green Monkey kidney cells were selected for  
 CC further analysis. A 7.4kb restriction map (about 9% of the HAV  
 CC genome) was constructed from 5 overlapping inserts. The sequence of  
 CC the first 3.3kb (approx.) from the 5'-terminus was determined. An  
 CC amino acid sequence was deduced from the entire clone and an open  
 CC reading frame was identified starting at position 238. A comparison  
 CC of the predicted HAV amino acid sequences with the known capsid  
 CC protein sequences of other picornaviruses (poliovirus, foot and  
 CC mouth disease virus and encephalomyelitis virus) revealed areas of  
 CC local homology.  
 CC (Note: Revised entry submitted to correct the patent number format of  
 CC US Government-owned NIS applications to prevent clashes with ongoing US  
 CC granted patent numbers. For further information please visit the Derwent



us-09-171-432a-39.rag

```

PS Claim 1; Fig 1; 18pp; English.
XX
CC The attenuated HAV is useful for inducing protective immunity against
CC HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
CC several nucleotide changes distributed throughout the genome, is
CC attenuated for chimpanzees, elicits serum neutralising antibodies, and is
CC suitable for use as an HAV vaccine. It is noted that not all the changes
CC are necessary for attenuation and use as a vaccine.
CC ('Updated on 25-MAR-2003 to correct PA field.')
CC ('Updated on 25-MAR-2003 to correct PI field.')
XX
SQ Sequence 2227 AA;
XX
Query Match 100.0%; Score 107; DB 11; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY 1 SHIECRKPYKRLKLVGVKOR 20
Db 792 SHIECRKPYKRLKLVGVKOR 811
RESULT 10
AAM34074
ID AAM34074 standard; Protein: 2227 AA.
XX
XX AAM34074:
AC 27-APR-1998 (first entry)
XX
DE Hepatitis A virus HM-175 protein sequence.
XX
XX HAV; attenuation: 2C protein; 2C chimeric hepatitis A virus;
KM infection; vaccine.
XX
XX Hepatitis A virus HM-175.
XX
XX Key Location/Qualifiers
FH 1..23
FT Protein /label= VP4
FT 24..245 /label= VP2
FT 246..491 /label= VP3
FT 492..791 /label= VP1
FT Protein 792..980 /label= 2A
FT Protein 981..1087 /label= 2B
FT Protein 1088..1422 /label= 2C
FT Protein 1423..1496 /label= 3A
FT Protein 1497..1519 /label= 3B
FT Protein 1520..1738 /label= 3C
FT Protein 1739..2227 /label= 3D
FT Protein
XX
XX MO9740166-A2.
XX
XX 30-OCT-1997.
XX
XX 18-APR-1997; 97WO-US06506.
XX
XX 19-APR-1996; 96US-0015642.
XX
XX (USSH ) US SEC DEPT HEALTH.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Emerson SU, Purcell RH, Raychaudhuri G;

```

XX WPI: 1997-535850/49.  
DR N-PSDB; AAT93023.  
XX  
PT Human attenuated HAV genome containing simian HAV 2C gene - useful  
PT as vaccines against HAV infection  
XX  
PS Disclosure; Fig 13A-D; 66pp; English.  
XX  
CC This protein sequence is encoded by the human hepatitis A virus  
CC (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain  
CC HAV/7 is obtained by passage of HM-175 in African Green Monkey  
CC kidney cells. A claimed DNA construct (I) comprises a genome of  
CC HAV, where the genome is a human attenuated HAV genome in which a  
CC region of the 2C gene has been replaced by a corresponding region  
CC from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The  
CC region of the 2C gene from AGM-27 contained in the construct  
CC preferably encodes amino acids 120-328 of the 2C protein, amino  
CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA  
CC transcript of (1); (2) a cell transfected with (1) or the RNA  
CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the  
CC HAV of (3); and (5) a host cell containing the HAV of (3). (1) or  
CC its RNA transcript, can be used as a vaccine for preventing HAV in  
CC a mammal. (1) or the RNA transcript can also be used to stimulate  
CC the production of protective antibodies in the mammal.  
XX  
SQ Sequence 2227 AA;  
Query Match 100.0%; Score 107; DB 18; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1,4e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SHIECRPYKELRLEVGKOR 20  
DB 792 SHIECRPYKELRLEVGKOR 811  
RESULT 11  
AAB18607  
ID AAB18607 standard; Protein: 2227 AA.  
XX  
AC AAB18607;  
XX  
DT 15-JAN-2001 (first entry)  
XX  
DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.  
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.  
XX Hepatitis A virus.  
XX  
PN US6113912-A.  
XX  
PD 05-SEP-2000.  
XX  
PF 07-JUN-1995; 95US-0475886.  
XX  
PR 18-SEP-1992; 92US-0947338.  
PR 17-SEP-1993; 93WO-US08610.  
PR 10-MAR-1995; 95US-0397232.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;  
XX  
DR WPI: 2000-586464/55.  
DR N-PSDB; AAA75476.  
XX  
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell  
PT line useful as vaccine for protecting humans against hepatitis A virus  
PT infection, has modified genome compared to wild type  
XX  
PS Disclosure; Fig 6A-K; 72pp; English.

XX  
CC The present sequence is derived from a wild type hepatitis A virus  
CC (HAV) strain HM-174. The sequence is modified to produce HAV which  
CC are adapted to growth in the human fibroblast-like cell line MRC-5.  
CC The HAV is able to propagate in MRC-5 cells and retain appropriate  
CC attenuation. It is useful as a live vaccine for prophylaxis of  
CC hepatitis A in humans and other primates.  
XX  
SQ Sequence 2227 AA;  
Query Match 100.0%; Score 107; DB 21; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1,4e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SHIECRPYKELRLEVGKOR 20  
DB 792 SHIECRPYKELRLEVGKOR 811  
RESULT 12  
AAB18608  
ID AAB18608 standard; Protein: 2227 AA.  
XX  
AC AAB18608;  
XX  
DT 15-JAN-2001 (first entry)  
XX  
DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.  
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;  
XX P-35 virus.  
XX  
OS Hepatitis A virus.  
XX  
PN US6113912-A.  
XX  
PD 05-SEP-2000.  
XX  
PF 07-JUN-1995; 95US-0475886.  
XX  
PR 18-SEP-1992; 92US-0947338.  
PR 17-SEP-1993; 93WO-US08610.  
PR 10-MAR-1995; 95US-0397232.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;  
XX  
DR WPI: 2000-586464/55.  
DR N-PSDB; AAA75477.  
XX  
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell  
PT line useful as vaccine for protecting humans against hepatitis A virus  
PT infection, has modified genome compared to wild type  
XX  
PS Disclosure; Columns 67-78; 72pp; English.  
XX  
CC The present sequence is derived from passage 35 of a wild type  
CC hepatitis A virus (HAV) strain HM-174. The resulting virus is  
CC designated P-35 virus. The sequence is modified to produce HAV which  
CC are adapted to growth in the human fibroblast-like cell line MRC-5.  
CC The HAV is able to propagate in MRC-5 cells and retain appropriate  
CC attenuation. It is useful as a live vaccine for prophylaxis of  
CC hepatitis A in humans and other primates.  
XX  
SQ Sequence 2227 AA;  
Query Match 100.0%; Score 107; DB 21; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1,4e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SHIECRPYKELRLEVGKOR 20  
DB 792 SHIECRPYKELRLEVGKOR 811

DB 792 SHIECRKPYKELRLEVGROR 811

RESULT 13  
AAB18609  
ID AAB18609 standard; Protein: 2227 AA.  
XX  
AC AAB18609;  
XX  
DT 15-JAN-2001 (first entry)  
XX  
DE Amino acid sequence of live attenuated Hepatitis A virus 4380.  
XX  
KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;  
KM HAV 4380.  
XX  
OS Hepatitis A virus.  
XX  
PN US6113912-A.  
XX  
PF 05-SEP-2000.  
XX  
PR 07-JUN-1995; 95US-0475886.  
XX  
PR 18-SEP-1992; 92US-0947338.  
PR 17-SEP-1993; 93WO-US08610.  
PR 10-MAR-1995; 95US-0397232.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;  
XX  
DR WPI: 2000-586464/55.  
XX  
N-PSDB: AAA75478.  
XX  
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell  
XX  
PT line useful as vaccine for protecting humans against hepatitis A virus  
XX  
PT infection, has modified genome compared to wild type  
XX  
PS Disclosure; Columns 93-104; 72pp; English.  
XX  
CC The present sequence is derived from a live attenuated hepatitis A  
XX  
CC virus (HAV) of the invention, designated HAV 4380. The sequence is  
XX  
CC produced by modifying wild type HAV strain HM-174. The HAV of the  
XX  
CC invention are adapted to growth in the human fibroblast-like cell  
XX  
CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain  
XX  
CC appropriate attenuation. It is useful as a live vaccine for prophylaxis  
XX  
CC of hepatitis A in humans and other primates.

Sequence 2227 AA:  
Query Match 100.0%; Score 107; DB 21; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGROR 20  
DB 792 SHIECRKPYKELRLEVGROR 811  
|||||

RESULT 14  
ABG31727  
ID ABG31727 standard; Protein: 2227 AA.  
XX  
AC ABG31727;  
XX  
DT 29-NOV-2002 (first entry)  
XX  
DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.  
XX  
KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.  
XX  
OS Hepatitis A virus strain HM-175.

XX US6423318-B1.  
PN  
XX  
PD 23-JUL-2002.  
XX  
PF 31-AUG-2000; 2000US-0653499.  
XX  
PR 07-JUN-1995; 95US-0475886.  
PR 17-SEP-1993; 93US-0397232.  
PR 17-SEP-1993; 93WO-US08610.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA (SMIR ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX  
PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;  
XX  
DR WPI: 2002-680946/73.  
XX  
N-PSDB: ABS52787.  
XX  
PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth  
XX  
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection  
XX  
PS Disclosure; Fig 6; 71pp; English.  
XX  
CC The invention relates to a polynucleotide which encodes a hepatitis A  
XX  
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell  
XX  
CC line). The polynucleotide is useful for preparing a vaccine against  
XX  
CC hepatitis A virus infection. This sequence represents a hepatitis A virus  
XX  
CC strain HM-175 polypeptide.

SQ Sequence 2227 AA:  
Query Match 100.0%; Score 107; DB 23; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGROR 20  
DB 792 SHIECRKPYKELRLEVGROR 811  
|||||

RESULT 15  
ABG31728  
ID ABG31728 standard; Protein: 2227 AA.  
XX  
AC ABG31728;  
XX  
DT 29-NOV-2002 (first entry)  
XX  
DE Hepatitis A virus mutant strain HM-175/7 (PHAV/7) polypeptide.  
XX  
KW Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;  
KM virucide; mutant; PHAV/7; mutlein.  
XX  
OS Hepatitis A virus strain HM-175.  
XX  
OS Synthetic.  
FH  
FT Key Location/Qualifiers  
FT MISC-difference 963  
FT MISC-difference 764 /label= Wild-type Lys substituted by Arg  
FT MISC-difference 764 /note= "Wild-type Glu substituted by Val"  
FT MISC-difference 821 /note= "Wild-type Asn substituted by Ser"  
FT MISC-difference 1052 /note= "Wild-type Ala substituted by Val"  
FT MISC-difference 1062 /note= "Wild-type Gly substituted by Ala"  
FT MISC-difference 1118 /note= "Wild-type Lys substituted by Met"  
FT MISC-difference 1151 /note= "Wild-type Glu substituted by Lys"  
FT MISC-difference 1163

```

FT      Misc-difference 1277 /note- "Wild-type Phe substituted by Ser"
FT      /note- "Wild-type Val substituted by Ile"
FT      Misc-difference 1500 /note- "Wild-type His substituted by Tyr"
FT      /note- "Wild-type His substituted by Tyr"
FT      Misc-difference 1805 /note- "Wild-type Asp substituted by Asn"
FT      /note- "Wild-type Asp substituted by Asn"
FT      Misc-difference 1930 /note- "Wild-type Ser substituted by Thr"
FT      /note- "Wild-type Ser substituted by Thr"
XX
XX      US6423318-B1.
XX
XX      23-JUL-2002.
XX
XX      31-AUG-2000; 2000US-0653499.
XX
XX      07-JUN-1995; 95US-0475886.
XX      17-SEP-1993; 93US-0397232.
XX      17-SEP-1993; 93WO-US08610.
PR
PR      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PR      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX      Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;
XX
XX      WPI: 2002-680946/73.
XX      N-PSDB; ABS52788.
DR
DR      New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT      in MRC-5 cells, useful for preparing a vaccine against HAV infection -
PT
XX      Example 3; Column 67-78; 71pp; English.
XX
XX      The invention relates to a polynucleotide which encodes a hepatitis A
CC      virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC      line). The polynucleotide is useful for preparing a vaccine against
CC      hepatitis A virus infection. This sequence represents a hepatitis A virus
CC      mutant strain HM-175/7 (pHAV/7) polypeptide.
XX
XX      Sequence 2227 AA;
SQ
SQ      Query Match 100.0%; Score 107; DB 23; Length 2227;
SQ      Best Local Similarity 100.0%; Pred. No. 1.4e-07;
SQ      Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 SHIECRKPYKELRLEVGR 20
QY      ||||||||||||||||
QY      792 SHIECRKPYKELRLEVGR 811

```

Search completed: October 1, 2003, 09:56:40  
 Job time : 45.5294 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 12.8235 seconds  
(without alignments)  
149,988 Million cell updates/sec

Title: US-09-171-432a-39  
Perfect score: 107  
Sequence: 1 SHIECRKPKYKELRLEVGKQR 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
al number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pIR1:\*  
2: pIR2:\*  
3: pIR3:\*  
4: pIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	56	2 P00431	genome polyprotein
2	107	100.0	341	2 S04137	genome polyprotein
3	107	100.0	852	1 GNNYHA	genome polyprotein
4	107	100.0	1358	2 A03905	genome polyprotein
5	107	100.0	2227	1 GNNYHM	genome polyprotein
6	107	100.0	2227	1 GNNYHR	genome polyprotein
7	107	100.0	2227	1 GNNYMK	genome polyprotein
8	107	100.0	2227	1 GNNYHB	genome polyprotein
9	105	98.1	56	2 P00427	genome polyprotein
10	104	97.2	56	2 P00429	genome polyprotein
11	102	95.3	319	2 JH0135	genome polyprotein
12	100	93.5	56	2 P00430	genome polyprotein
13	97	90.7	56	2 P00432	genome polyprotein
14	96	89.7	56	2 P00428	genome polyprotein
15	95	88.8	56	2 P00434	genome polyprotein
16	87	81.3	55	2 P00433	genome polyprotein
17	86	80.4	56	2 P00436	genome polyprotein
18	86	80.4	2230	1 GNNYSA	genome polyprotein
19	84.5	79.0	55	2 P00435	genome polyprotein
20	84.5	79.0	839	1 GNNYS2	genome polyprotein
21	81	75.7	56	2 P00437	genome polyprotein
22	46	43.0	420	2 S62541	hypothetical prote
23	44	41.1	374	2 C84040	hypothetical prote
24	43.5	40.7	1663	1 C3RT	complement C3 prec
25	43	40.2	332	2 P96568	probable lipase, 2
26	43	40.2	339	2 T18926	hypothetical prote
27	43	40.2	343	2 AHI823	30S ribosomal prot
28	43	40.2	346	2 T27896	hypothetical prote
29	43	40.2	391	1 VHNZ3	nucleocapsid prote

30	43	40.2	554	2 T49917	hypothetical prote
31	42.5	39.7	140	2 E64364	hypothetical prote
32	42	39.3	121	2 T17708	hypothetical prote
33	42	39.3	214	2 E82679	chaperone XP1452 l
34	42	39.3	286	2 S07532	putf II/9-1 protei
35	42	39.3	329	2 T25067	hypothetical prote
36	42	39.3	494	2 S62902	legumin 2 precursor
37	42	39.3	723	2 D71091	methionine-tRNA 11
38	42	39.3	861	2 B84963	DNA topoisomerase
39	42	39.3	1087	2 T16876	hypothetical prote
40	42	39.3	1157	2 T43259	pyruvate (flavodox
41	42	39.3	1270	2 T22615	hypothetical prote
42	41.5	38.8	286	2 H87506	hypothetical prote
43	41.5	38.8	606	2 A12649	phosphogluconate d
44	41.5	38.8	606	2 H97431	phosphogluconate d
45	41	38.3	144	2 H96511	AP2 domain contain

#### ALIGNMENTS

RESULT 1  
P00431  
genome polyprotein - human hepatitis A virus (strain No. 4) (fragment)  
C:Species: human hepatitis A virus  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: P00431  
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totstuka, A.; Naiman, O.V.; Siegl, G.; Wl  
J. Gen. Virol. 73, 1365-1377, 1992  
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge  
A:Reference number: P00427; MUID:92300330; PMID:1318940  
A:Accession: P00431  
A:Molecule type: mRNA  
A:Residues: 1-56 <ROB>  
A>Note: this protein is from the VP1/2A Junction region  
C:Superfamily: hepatitis A virus genome polyprotein  
C:Keywords: coat protein; core protein; polyprotein

Query Match 100.0%; Score 107; DB 2; Length 56;  
Best local similarity 100.0%; Pred. No. 3.9e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPKYKELRLEVGKQR 20  
DB 29 SHIECRKPKYKELRLEVGKQR 48  
|||||  
RESULT 2  
S04137  
genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)  
C:Species: human hepatitis A virus  
C>Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 21-Jul-2000  
C:Accession: S04137  
R:Andonov, A.P.; Lau, P.C.K.; Chaudhary, R.  
Nucleic Acids Res. 17, 3594, 1989  
A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A vir  
A:Reference number: S04137; MUID:89263805; PMID:2542903  
A:Accession: S04137  
A:Molecule type: mRNA  
A:Residues: 1-341 <AND>  
A:Cross-references: EMBL:X14666; NID:g62301; PIDN:CAA32794.1; PID:g4377576  
C:Genetics:  
A:Gene: VP1  
C:Superfamily: hepatitis A virus genome polyprotein  
C:Keywords: coat protein; polyprotein  
F:2-340/Product: coat protein ID (VP1) #status predicted <ANT>  
Query Match 100.0%; Score 107; DB 2; Length 341;  
Best local similarity 100.0%; Pred. No. 2.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SHIECRKPKYKELRLEVGKQR 20  
|||||

Db 296 SHIECRPKYKELRLEVGKOR 315

## RESULT 3

GNNYHM

genome polyprotein - human hepatitis A virus (strain CM326) (fragment)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A  
C:Species: human hepatitis A virus  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Jul-1999  
C:Accession: A03904  
R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.W.  
J. Virol. 54, 247-255, 1985  
A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.  
A:Reference number: A03904; MUID:85185648; PMID:2985793  
A:Accession: A03904  
A:Molecule type: genomic RNA  
A:Residues: 1-852 <LIN>  
A:Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593  
Superfamily: hepatitis A virus genome polyprotein  
C:Keywords: coat protein; core protein; polyprotein  
F:1-245/Product: coat protein 1A #status predicted <C1A>  
F:246-491/Product: coat protein 1B #status predicted <C1B>  
F:492-836/Product: coat protein 1C #status predicted <C1C>  
F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 100.0%; Score 107; DB 1; Length 852;  
Best Local Similarity 100.0%; Pred. No. 6.3e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRPKYKELRLEVGKOR 20

Db 792 SHIECRPKYKELRLEVGKOR 811

## RESULT 4

GNNYHM

genome polyprotein (version 2) - human hepatitis A virus (fragments)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; protein  
C:Species: human hepatitis A virus  
C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 15-Nov-1996  
C:Accession: A03905  
R:Baroudy, B.M.; Ticehurst, J.R.; Mele, T.A.; Maizel Jr., J.V.; Purcell, R.H.; Feinstor  
Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985  
A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA  
A:Reference number: A03905; MUID:85186289; PMID:2984684  
A:Accession: A03905  
A:Molecule type: genomic RNA  
A:Residues: 1-1358 <BAR>  
Superfamily: hepatitis A virus genome polyprotein  
C:Keywords: coat protein; core protein; polyprotein  
F:1-245/Product: coat protein 1A #status predicted <C1A>  
F:246-491/Product: coat protein 1B #status predicted <C1B>  
F:492-836/Product: coat protein 1C #status predicted <C1C>  
F:837-854/Product: core protein 2A (fragment) #status predicted <C2A>  
F:855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match 100.0%; Score 107; DB 2; Length 1358;  
Best Local Similarity 100.0%; Pred. No. 1e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRPKYKELRLEVGKOR 20

Db 792 SHIECRPKYKELRLEVGKOR 811

## RESULT 5

GNNYHM

genome polyprotein - human hepatitis A virus (strain HM-175, wild type)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro  
C:Species: human hepatitis A virus  
A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 16-Jul-1999

C:Accession: A25981

R:Chen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with

A:Reference number: A25981; MUID:87061253; PMID:3023706

A:Accession: A25981

A:Molecule type: genomic RNA

A:Residues: 1-2227 <COH>

A:Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-23/Product: coat protein 1A #status predicted <VP1>

F:24-243/Product: coat protein 1B #status predicted <VP2>

F:246-491/Product: coat protein 1C #status predicted <VP3>

F:492-791/Product: coat protein 1D #status predicted <VP4>

F:792-980/Product: core protein 2A #status predicted <C2A>

F:981-1087/Product: core protein 2B #status predicted <C2B>

F:1088-1422/Product: core protein 2C #status predicted <C2C>

F:1423-1496/Product: protein 3A #status predicted <C3A>

F:1497-1519/Product: protein 3B #status predicted <C3B>

F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 107; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRPKYKELRLEVGKOR 20

Db 792 SHIECRPKYKELRLEVGKOR 811

## RESULT 6

GNNYHM

genome polyprotein - human hepatitis A virus  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core  
NA polymerase (EC 2.7.7.48), protein 3D  
C:Species: human hepatitis A virus  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Jul-1999  
C:Accession: A03903  
R:Naftarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N  
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985  
A:Title: Primary structure and gene organization of human hepatitis A virus.  
A:Reference number: A03903; MUID:85190549; PMID:2986127  
A:Accession: A03903  
A:Molecule type: genomic RNA  
A:Residues: 1-2227 <NAJ>  
A:Cross-references: GB:K02990; NID:g329596; PIDN:AAA45472.1; PID:g329597  
C:Superfamily: hepatitis A virus genome polyprotein  
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr  
F:1-245/Product: coat protein 1A #status predicted <C1A>  
F:246-491/Product: coat protein 1B #status predicted <C1B>  
F:492-836/Product: coat protein 1C #status predicted <C1C>  
F:837-980/Product: coat protein 2A #status predicted <C2A>  
F:981-1076/Product: core protein 2B #status predicted <C2B>  
F:1077-1422/Product: core protein 2C #status predicted <C2C>  
F:1423-1496/Product: protein 3A #status predicted <C3A>  
F:1497-1519/Product: protein 3B #status predicted <C3B>  
F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>  
F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 107; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRPKYKELRLEVGKOR 20

Db 792 SHIECRPKYKELRLEVGKOR 811

## RESULT 7

GNNYMK genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core protein 2B  
NA polymerase (EC 2.7.7.48), protein 3D  
C:Species: human hepatitis A virus  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1968 #sequence\_revision 30-Jun-1988 #text\_change 16-Jul-1999  
A:Accession: A94149; A25914; A94508  
R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Puccell, R.  
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987  
A>Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with  
A:Reference number: A94149; MUID:87175701; PMID:3031666  
A:Accession: A94149  
A>Status: nucleic acid sequence not shown  
A:Molecule type: genomic RNA  
A:Residues: 1-2227 <CO>  
A:Cross-references: EMBL:M1632; NID:g329594; PID:AAA45471.1; PID:g329595  
A:Note: submitted to Genbank, August 1987  
C:Superfamily: hepatitis A virus genome polyprotein  
E:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans  
F:146-491/Product: coat protein 1A #status predicted <P1A>  
F:149-836/Product: coat protein 1B #status predicted <P1B>  
F:149-836/Product: coat protein 1C #status predicted <P1C>  
F:149-836/Product: coat protein 2A #status predicted <P2A>  
F:149-836/Product: coat protein 2B #status predicted <P2B>  
F:149-836/Product: coat protein 2C #status predicted <P2C>  
F:149-836/Product: coat protein 3A #status predicted <P3A>  
F:149-836/Product: coat protein 3B #status predicted <P3B>  
F:149-836/Product: coat protein 3C #status predicted <P3C>  
F:149-836/Product: coat protein 3D #status predicted <P3D>

Query Match 100.0%; Score 107; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.7e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Ox 1 SHICRRPYKELRVGKOR 20  
Db 792 SHICRRPYKELRVGKOR 811

RESULT 8  
GNNYMK genome polyprotein - human hepatitis A virus (strain MBB)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro  
VPG; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D  
C:Species: human hepatitis A virus  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 10-May-1996  
A:Accession: JS0303  
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat  
us Res. 8, 153-171, 1987  
A:Reference number: JS0303; MUID:68045071; PMID:2823500  
A:Accession: JS0303  
A:Molecule type: genomic RNA  
A:Residues: 1-2227 <PAU>  
A:Cross-references: EMBL:M20273  
C:Superfamily: hepatitis A virus genome polyprotein  
E:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydi  
F:1-23/Product: coat protein 1A #status predicted <VP1>  
F:24-246/Product: coat protein 1B #status predicted <VP2>  
F:247-491/Product: coat protein 1C #status predicted <VP3>  
F:492-836/Product: coat protein 1D #status predicted <VP4>  
F:837-980/Product: core protein 2A #status predicted <P2A>  
F:981-1108/Product: core protein 2B #status predicted <P2B>  
F:1109-1438/Product: core protein 2C #status predicted <P2C>  
F:1439-1496/Product: core protein 3A #status predicted <P3A>  
F:1497-1519/Product: genome-linked protein VPG #status predicted <VPG>  
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>  
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 107; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.7e-08;

```

Matches 20: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPYKELRLEVCGKOR 20
    |||||||
Db 792 SHIECRKPYKELRLEVCGKOR 811

RESULT 9
P00427
genome polypotein - human hepatitis A virus (strain BP-35.730) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: P00427
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Wl
J:Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A:Reference number: P00427; MUID:92300330; PMID:1318940
A:Accession: P00427
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
C:Comment: This protein is from the VP1/2A Junction region.
C:Superfamily: hepatitis A virus genome polypotein
C:Keywords: coat protein; core protein; polypotein

Query Match 98.1%; Score 105; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 8.2e-10;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPYKELRLEVCGKOR 20
    |||||||
Db 29 SHIECRKPYKELRLEVCGKOR 48

RESULT 10
P00429
genome polypotein - human hepatitis A virus (strain PRCL6) (fragment)
C:Species: human hepatitis A virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: P00429
R:Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Wl
J:Gen. Virol. 73, 1365-1377, 1992
A:Title: Genetic relatedness of hepatitis A virus strains recovered from different ge
A:Reference number: P00427; MUID:92300330; PMID:1318940
A:Accession: P00429
A:Molecule type: mRNA
A:Residues: 1-56 <ROB>
A>Note: this protein is from the VP1/2A Junction region
C:Superfamily: hepatitis A virus genome polypotein
C:Keywords: coat protein; core protein; polypotein

Query Match 97.2%; Score 104; DB 2; Length 56;
Best Local Similarity 95.0%; Pred. No. 1.2e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SHIECRKPYKELRLEVCGKOR 20
    |||||||
Db 29 SHIECRKPYKELRLEVCGKOR 48

RESULT 11
JH0135
genome polypotein - human hepatitis A virus (strain MS-1) (fragment)
M:Contains: amino end of core protein 2A; coat protein 1D
C:Species: human hepatitis A virus
A>Note: host Homo Sapiens (man)
C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 08-Apr-1994
C:Accession: JH0135
R:Robertson, B.H.; Brown, V.K.; Bradley, D.W.
Virus Res. 8, 309-316, 1987
A:Title: Nucleic acid sequence of the VP1 region of attenuated MS-1 hepatitis A virus
A:Reference number: JH0135; MUID:88129044; PMID:2829458
A:Accession: JH0135
A:Molecule type: genomic RNA

```

A;Residues: 1-319 <ROB>  
A;Cross-references: GB:M22821  
C;Superfamily: hepatitis A virus genome polyprotein  
C;Keywords: coat protein; core protein; glycoprotein; polyprotein  
F:1-300/Product: coat protein ID #status predicted <CDB>  
F:301-319/Product: core protein 2A (fragment) #status predicted <C2A>  
F:237/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 95.3%; Score 102; DB 2; Length 319;  
Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLVGVKOR 19  
|||||  
Db 301 SHIECRKPYKELRLVGVKQ 319

## RESULT 12

genome polyprotein - human hepatitis A virus (strain S23-1) (fragment)  
P00430  
C;Accession: P00430  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999

R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992  
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr  
A;Reference number: P00427; MUID:92300330; PMID:1318940  
A;Accession: P00430  
A;Molecule type: mRNA  
A;Residues: 1-56 <ROB>  
A;Note: this protein is from the VP1/2A Junction region  
C;Superfamily: hepatitis A virus genome polyprotein  
C;Keywords: coat protein; core protein; polyprotein

Query Match 93.5%; Score 100; DB 2; Length 56;  
Best Local Similarity 95.0%; Pred. No. 5.2e-09;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLVGVKOR 20  
|||||  
Db 29 SHIECRKPYKELRLVGVKOR 48

## RESULT 13

genome polyprotein - human hepatitis A virus (strain CF-53) (fragment)  
P00432  
C;Species: human hepatitis A virus  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999

R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992  
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different geogr  
A;Reference number: P00427; MUID:92300330; PMID:1318940  
A;Accession: P00432  
A;Molecule type: mRNA  
A;Residues: 1-56 <ROB>  
A;Note: this protein is from the VP1/2A Junction region  
C;Superfamily: hepatitis A virus genome polyprotein  
C;Keywords: coat protein; core protein; polyprotein

Query Match 90.7%; Score 97; DB 2; Length 56;  
Best Local Similarity 95.0%; Pred. No. 1.6e-08;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLVGVKOR 20  
|||||  
Db 29 SHIECRKPYKELRLVGVKOR 48

## RESULT 14

genome polyprotein - human hepatitis A virus (strain TKM002) (fragment)  
P00428  
C;Species: human hepatitis A virus

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C;Accession: P00428  
R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992  
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different ge  
A;Reference number: P00427; MUID:92300330; PMID:1318940  
A;Accession: P00428  
A;Molecule type: mRNA  
A;Residues: 1-56 <ROB>  
A;Note: this protein is from the VP1/2A Junction region  
C;Superfamily: hepatitis A virus genome polyprotein  
C;Keywords: coat protein; core protein; polyprotein

Query Match 89.7%; Score 96; DB 2; Length 56;  
Best Local Similarity 95.0%; Pred. No. 2.3e-08;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLVGVKOR 20  
|||||  
Db 29 SHIECRKPYKELRLVGVKOR 48

## RESULT 15

genome polyprotein - human hepatitis A virus (strain KPH) (fragment)  
P00434  
C;Species: human hepatitis A virus  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999

R;Robertson, B.H.; Jansen, R.W.; Khanna, B.; Totsuka, A.; Nainan, O.V.; Siegl, G.; Widel, J. Gen. Virol. 73, 1365-1377, 1992  
A;Title: Genetic relatedness of hepatitis A virus strains recovered from different ge  
A;Reference number: P00427; MUID:92300330; PMID:1318940  
A;Accession: P00434  
A;Molecule type: mRNA  
A;Residues: 1-56 <ROB>  
A;Note: this protein is from the VP1/2A Junction region  
C;Superfamily: hepatitis A virus genome polyprotein  
C;Keywords: coat protein; core protein; polyprotein

Query Match 88.8%; Score 95; DB 2; Length 56;  
Best Local Similarity 95.0%; Pred. No. 3.3e-08;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLVGVKOR 20  
|||||  
Db 29 SHIECRKPYKELRLVGVKOR 48

Search completed: October 1, 2003, 10:04:33  
Job time : 15.8235 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 ; Search time 6.58824 Seconds

(without alignments)  
142.760 Million cell updates/sec

Title: US-09-171-432A-39

Perfect score: 107

Sequence: 1 SHIECRKPYKELRELVGKOR 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

al number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot\_41:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	107	100.0	341 1	P03672 hepatitis a
2	107	100.0	852 1	P06442 hepatitis a
3	107	100.0	2226 1	P26580 hepatitis a
4	107	100.0	2226 1	P26581 hepatitis a
5	107	100.0	2226 1	P26582 hepatitis a
6	107	100.0	2227 1	P08617 hepatitis a
7	107	100.0	2227 1	P08617 hepatitis a
8	107	100.0	2227 1	P08617 hepatitis a
9	95	88.8	808 1	P13901 hepatitis a
10	86	80.4	2230 1	P02381 hepatitis a
11	84.5	79.0	839 1	P03178 simian hepa
12	46	43.0	420 1	P14553 simian hepa
13	45	42.1	1819 1	P09877 schistosom
14	43.5	40.7	1663 1	P01026 rattus norv
15	43	40.2	391 1	P24566 human resp
16	42	39.3	210 1	P09d67 xyella fas
17	42	39.3	286 1	P22311 sclera copr
18	42	39.3	723 1	P058721 pyrococcus
19	42	39.3	861 1	P57371 buchnera ap
20	41	38.3	273 1	P08493 buchnera ap
21	41	38.3	298 1	P08493 buchnera ap
22	41	38.3	319 1	U087399 synecococc
23	41	38.3	364 1	SUCC_METVA
24	41	38.3	370 1	T2M2_METVA
25	41	38.3	391 1	NCAP_BR5V3
26	41	38.3	391 1	NCAP_BR5V3
27	41	38.3	391 1	NCAP_BR5V3
28	41	38.3	391 1	NCAP_BR5V3
29	41	38.3	492 1	ANKH_HUMAN
30	41	38.3	492 1	ANKH_MOUSE
31	41	38.3	492 1	ANKH_MOUSE
32	41	38.3	492 1	ANKH_MOUSE
33	41	38.3	840 1	NI61_HUMAN

34	41	38.3	843 1	NI61_MOUSE	Q09K10 mus musculu
35	41	38.3	843 1	NI61_MOUSE	Q62765 rattus norv
36	41	38.3	955 1	T150_HUMAN	O9Y2W1 homo sapien
37	41	38.3	4543 1	LRP1_CHICK	P98157 gallus galli
38	41	38.3	4544 1	LRP1_HUMAN	O07954 homo sapien
39	40.5	37.9	141 1	V16K_TRYPC	P05076 tobacco rat
40	40.5	37.9	274 1	NAGB_FUSNN	O8REG1 fusobacteri
41	40.5	37.9	1120 1	RPOW_SCHPO	O13993 schizosacch
42	40.5	37.9	1663 1	CO3_MOUSE	P01027 mus musculu
43	40.5	37.9	2198 1	YLJ2_CAEEL	P34367 caenorhabdi
44	40	37.4	69 1	MOP_HAEIN	P45183 haemophilus
45	40	37.4	238 1	PTH_AGRTS	Q8UD97 agrobacteri

ALIGNMENTS

```
RESULT 1
POLG_HPAV1 STANDARD: PRT; 341 AA.
AC P13672:
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP3; Core protein
DE P2A] (Fragment).
OS Hepatitis A virus (strain LCD-1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263805; PubMed=2542903;
RA Andonov A.P., Lau P., Chaudhary R.;
RT "Nucleotide sequence of the VP1 gene from a Chinese strain of
RT hepatitis A virus (HAV).";
RL Nucleic Acids Res. 17:3594-3594(1989).
CC -i- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -i- PTM: SPECIFIC ENZYMOLOGIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14666; GAA32794.1; -.
DR PIR: S04137; S04137.
KM POLYPROTEIN: Coat protein; Core protein.
FT NON_TER 1
FT CHAIN 1
FT CHAIN 2
FT CHAIN 3
FT CHAIN 4
FT CHAIN 5
FT CHAIN 6
FT CHAIN 7
FT CHAIN 8
FT CHAIN 9
FT CHAIN 10
FT CHAIN 11
FT CHAIN 12
FT CHAIN 13
FT CHAIN 14
FT CHAIN 15
FT CHAIN 16
FT CHAIN 17
FT CHAIN 18
FT CHAIN 19
FT CHAIN 20
FT CHAIN 21
FT CHAIN 22
FT CHAIN 23
FT CHAIN 24
FT CHAIN 25
FT CHAIN 26
FT CHAIN 27
FT CHAIN 28
FT CHAIN 29
FT CHAIN 30
FT CHAIN 31
FT CHAIN 32
FT CHAIN 33
FT CHAIN 34
FT CHAIN 35
FT CHAIN 36
FT CHAIN 37
FT CHAIN 38
FT CHAIN 39
FT CHAIN 40
FT CHAIN 41
FT CHAIN 42
FT CHAIN 43
FT CHAIN 44
FT CHAIN 45
FT CHAIN 46
FT CHAIN 47
FT CHAIN 48
FT CHAIN 49
FT CHAIN 50
FT CHAIN 51
FT CHAIN 52
FT CHAIN 53
FT CHAIN 54
FT CHAIN 55
FT CHAIN 56
FT CHAIN 57
FT CHAIN 58
FT CHAIN 59
FT CHAIN 60
FT CHAIN 61
FT CHAIN 62
FT CHAIN 63
FT CHAIN 64
FT CHAIN 65
FT CHAIN 66
FT CHAIN 67
FT CHAIN 68
FT CHAIN 69
FT CHAIN 70
FT CHAIN 71
FT CHAIN 72
FT CHAIN 73
FT CHAIN 74
FT CHAIN 75
FT CHAIN 76
FT CHAIN 77
FT CHAIN 78
FT CHAIN 79
FT CHAIN 80
FT CHAIN 81
FT CHAIN 82
FT CHAIN 83
FT CHAIN 84
FT CHAIN 85
FT CHAIN 86
FT CHAIN 87
FT CHAIN 88
FT CHAIN 89
FT CHAIN 90
FT CHAIN 91
FT CHAIN 92
FT CHAIN 93
FT CHAIN 94
FT CHAIN 95
FT CHAIN 96
FT CHAIN 97
FT CHAIN 98
FT CHAIN 99
FT CHAIN 100
FT CHAIN 101
FT CHAIN 102
FT CHAIN 103
FT CHAIN 104
FT CHAIN 105
FT CHAIN 106
FT CHAIN 107
FT CHAIN 108
FT CHAIN 109
FT CHAIN 110
FT CHAIN 111
FT CHAIN 112
FT CHAIN 113
FT CHAIN 114
FT CHAIN 115
FT CHAIN 116
FT CHAIN 117
FT CHAIN 118
FT CHAIN 119
FT CHAIN 120
FT CHAIN 121
FT CHAIN 122
FT CHAIN 123
FT CHAIN 124
FT CHAIN 125
FT CHAIN 126
FT CHAIN 127
FT CHAIN 128
FT CHAIN 129
FT CHAIN 130
FT CHAIN 131
FT CHAIN 132
FT CHAIN 133
FT CHAIN 134
FT CHAIN 135
FT CHAIN 136
FT CHAIN 137
FT CHAIN 138
FT CHAIN 139
FT CHAIN 140
FT CHAIN 141
FT CHAIN 142
FT CHAIN 143
FT CHAIN 144
FT CHAIN 145
FT CHAIN 146
FT CHAIN 147
FT CHAIN 148
FT CHAIN 149
FT CHAIN 150
FT CHAIN 151
FT CHAIN 152
FT CHAIN 153
FT CHAIN 154
FT CHAIN 155
FT CHAIN 156
FT CHAIN 157
FT CHAIN 158
FT CHAIN 159
FT CHAIN 160
FT CHAIN 161
FT CHAIN 162
FT CHAIN 163
FT CHAIN 164
FT CHAIN 165
FT CHAIN 166
FT CHAIN 167
FT CHAIN 168
FT CHAIN 169
FT CHAIN 170
FT CHAIN 171
FT CHAIN 172
FT CHAIN 173
FT CHAIN 174
FT CHAIN 175
FT CHAIN 176
FT CHAIN 177
FT CHAIN 178
FT CHAIN 179
FT CHAIN 180
FT CHAIN 181
FT CHAIN 182
FT CHAIN 183
FT CHAIN 184
FT CHAIN 185
FT CHAIN 186
FT CHAIN 187
FT CHAIN 188
FT CHAIN 189
FT CHAIN 190
FT CHAIN 191
FT CHAIN 192
FT CHAIN 193
FT CHAIN 194
FT CHAIN 195
FT CHAIN 196
FT CHAIN 197
FT CHAIN 198
FT CHAIN 199
FT CHAIN 200
FT CHAIN 201
FT CHAIN 202
FT CHAIN 203
FT CHAIN 204
FT CHAIN 205
FT CHAIN 206
FT CHAIN 207
FT CHAIN 208
FT CHAIN 209
FT CHAIN 210
FT CHAIN 211
FT CHAIN 212
FT CHAIN 213
FT CHAIN 214
FT CHAIN 215
FT CHAIN 216
FT CHAIN 217
FT CHAIN 218
FT CHAIN 219
FT CHAIN 220
FT CHAIN 221
FT CHAIN 222
FT CHAIN 223
FT CHAIN 224
FT CHAIN 225
FT CHAIN 226
FT CHAIN 227
FT CHAIN 228
FT CHAIN 229
FT CHAIN 230
FT CHAIN 231
FT CHAIN 232
FT CHAIN 233
FT CHAIN 234
FT CHAIN 235
FT CHAIN 236
FT CHAIN 237
FT CHAIN 238
FT CHAIN 239
FT CHAIN 240
FT CHAIN 241
FT CHAIN 242
FT CHAIN 243
FT CHAIN 244
FT CHAIN 245
FT CHAIN 246
FT CHAIN 247
FT CHAIN 248
FT CHAIN 249
FT CHAIN 250
FT CHAIN 251
FT CHAIN 252
FT CHAIN 253
FT CHAIN 254
FT CHAIN 255
FT CHAIN 256
FT CHAIN 257
FT CHAIN 258
FT CHAIN 259
FT CHAIN 260
FT CHAIN 261
FT CHAIN 262
FT CHAIN 263
FT CHAIN 264
FT CHAIN 265
FT CHAIN 266
FT CHAIN 267
FT CHAIN 268
FT CHAIN 269
FT CHAIN 270
FT CHAIN 271
FT CHAIN 272
FT CHAIN 273
FT CHAIN 274
FT CHAIN 275
FT CHAIN 276
FT CHAIN 277
FT CHAIN 278
FT CHAIN 279
FT CHAIN 280
FT CHAIN 281
FT CHAIN 282
FT CHAIN 283
FT CHAIN 284
FT CHAIN 285
FT CHAIN 286
FT CHAIN 287
FT CHAIN 288
FT CHAIN 289
FT CHAIN 290
FT CHAIN 291
FT CHAIN 292
FT CHAIN 293
FT CHAIN 294
FT CHAIN 295
FT CHAIN 296
FT CHAIN 297
FT CHAIN 298
FT CHAIN 299
FT CHAIN 300
FT CHAIN 301
FT CHAIN 302
FT CHAIN 303
FT CHAIN 304
FT CHAIN 305
FT CHAIN 306
FT CHAIN 307
FT CHAIN 308
FT CHAIN 309
FT CHAIN 310
FT CHAIN 311
FT CHAIN 312
FT CHAIN 313
FT CHAIN 314
FT CHAIN 315
FT CHAIN 316
FT CHAIN 317
FT CHAIN 318
FT CHAIN 319
FT CHAIN 320
FT CHAIN 321
FT CHAIN 322
FT CHAIN 323
FT CHAIN 324
FT CHAIN 325
FT CHAIN 326
FT CHAIN 327
FT CHAIN 328
FT CHAIN 329
FT CHAIN 330
FT CHAIN 331
FT CHAIN 332
FT CHAIN 333
FT CHAIN 334
FT CHAIN 335
FT CHAIN 336
FT CHAIN 337
FT CHAIN 338
FT CHAIN 339
FT CHAIN 340
FT CHAIN 341
FT CHAIN 342
FT CHAIN 343
FT CHAIN 344
FT CHAIN 345
FT CHAIN 346
FT CHAIN 347
FT CHAIN 348
FT CHAIN 349
FT CHAIN 350
FT CHAIN 351
FT CHAIN 352
FT CHAIN 353
FT CHAIN 354
FT CHAIN 355
FT CHAIN 356
FT CHAIN 357
FT CHAIN 358
FT CHAIN 359
FT CHAIN 360
FT CHAIN 361
FT CHAIN 362
FT CHAIN 363
FT CHAIN 364
FT CHAIN 365
FT CHAIN 366
FT CHAIN 367
FT CHAIN 368
FT CHAIN 369
FT CHAIN 370
FT CHAIN 371
FT CHAIN 372
FT CHAIN 373
FT CHAIN 374
FT CHAIN 375
FT CHAIN 376
FT CHAIN 377
FT CHAIN 378
FT CHAIN 379
FT CHAIN 380
FT CHAIN 381
FT CHAIN 382
FT CHAIN 383
FT CHAIN 384
FT CHAIN 385
FT CHAIN 386
FT CHAIN 387
FT CHAIN 388
FT CHAIN 389
FT CHAIN 390
FT CHAIN 391
FT CHAIN 392
FT CHAIN 393
FT CHAIN 394
FT CHAIN 395
FT CHAIN 396
FT CHAIN 397
FT CHAIN 398
FT CHAIN 399
FT CHAIN 400
FT CHAIN 401
FT CHAIN 402
FT CHAIN 403
FT CHAIN 404
FT CHAIN 405
FT CHAIN 406
FT CHAIN 407
FT CHAIN 408
FT CHAIN 409
FT CHAIN 410
FT CHAIN 411
FT CHAIN 412
FT CHAIN 413
FT CHAIN 414
FT CHAIN 415
FT CHAIN 416
FT CHAIN 417
FT CHAIN 418
FT CHAIN 419
FT CHAIN 420
FT CHAIN 421
FT CHAIN 422
FT CHAIN 423
FT CHAIN 424
FT CHAIN 425
FT CHAIN 426
FT CHAIN 427
FT CHAIN 428
FT CHAIN 429
FT CHAIN 430
FT CHAIN 431
FT CHAIN 432
FT CHAIN 433
FT CHAIN 434
FT CHAIN 435
FT CHAIN 436
FT CHAIN 437
FT CHAIN 438
FT CHAIN 439
FT CHAIN 440
FT CHAIN 441
FT CHAIN 442
FT CHAIN 443
FT CHAIN 444
FT CHAIN 445
FT CHAIN 446
FT CHAIN 447
FT CHAIN 448
FT CHAIN 449
FT CHAIN 450
FT CHAIN 451
FT CHAIN 452
FT CHAIN 453
FT CHAIN 454
FT CHAIN 455
FT CHAIN 456
FT CHAIN 457
FT CHAIN 458
FT CHAIN 459
FT CHAIN 460
FT CHAIN 461
FT CHAIN 462
FT CHAIN 463
FT CHAIN 464
FT CHAIN 465
FT CHAIN 466
FT CHAIN 467
FT CHAIN 468
FT CHAIN 469
FT CHAIN 470
FT CHAIN 471
FT CHAIN 472
FT CHAIN 473
FT CHAIN 474
FT CHAIN 475
FT CHAIN 476
FT CHAIN 477
FT CHAIN 478
FT CHAIN 479
FT CHAIN 480
FT CHAIN 481
FT CHAIN 482
FT CHAIN 483
FT CHAIN 484
FT CHAIN 485
FT CHAIN 486
FT CHAIN 487
FT CHAIN 488
FT CHAIN 489
FT CHAIN 490
FT CHAIN 491
FT CHAIN 492
FT CHAIN 493
FT CHAIN 494
FT CHAIN 495
FT CHAIN 496
FT CHAIN 497
FT CHAIN 498
FT CHAIN 499
FT CHAIN 500
FT CHAIN 501
FT CHAIN 502
FT CHAIN 503
FT CHAIN 504
FT CHAIN 505
FT CHAIN 506
FT CHAIN 507
FT CHAIN 508
FT CHAIN 509
FT CHAIN 510
FT CHAIN 511
FT CHAIN 512
FT CHAIN 513
FT CHAIN 514
FT CHAIN 515
FT CHAIN 516
FT CHAIN 517
FT CHAIN 518
FT CHAIN 519
FT CHAIN 520
FT CHAIN 521
FT CHAIN 522
FT CHAIN 523
FT CHAIN 524
FT CHAIN 525
FT CHAIN 526
FT CHAIN 527
FT CHAIN 528
FT CHAIN 529
FT CHAIN 530
FT CHAIN 531
FT CHAIN 532
FT CHAIN 533
FT CHAIN 534
FT CHAIN 535
FT CHAIN 536
FT CHAIN 537
FT CHAIN 538
FT CHAIN 539
FT CHAIN 540
FT CHAIN 541
FT CHAIN 542
FT CHAIN 543
FT CHAIN 544
FT CHAIN 545
FT CHAIN 546
FT CHAIN 547
FT CHAIN 548
FT CHAIN 549
FT CHAIN 550
FT CHAIN 551
FT CHAIN 552
FT CHAIN 553
FT CHAIN 554
FT CHAIN 555
FT CHAIN 556
FT CHAIN 557
FT CHAIN 558
FT CHAIN 559
FT CHAIN 560
FT CHAIN 561
FT CHAIN 562
FT CHAIN 563
FT CHAIN 564
FT CHAIN 565
FT CHAIN 566
FT CHAIN 567
FT CHAIN 568
FT CHAIN 569
FT CHAIN 570
FT CHAIN 571
FT CHAIN 572
FT CHAIN 573
FT CHAIN 574
FT CHAIN 575
FT CHAIN 576
FT CHAIN 577
FT CHAIN 578
FT CHAIN 579
FT CHAIN 580
FT CHAIN 581
FT CHAIN 582
FT CHAIN 583
FT CHAIN 584
FT CHAIN 585
FT CHAIN 586
FT CHAIN 587
FT CHAIN 588
FT CHAIN 589
FT CHAIN 590
FT CHAIN 591
FT CHAIN 592
FT CHAIN 593
FT CHAIN 594
FT CHAIN 595
FT CHAIN 596
FT CHAIN 597
FT CHAIN 598
FT CHAIN 599
FT CHAIN 600
FT CHAIN 601
FT CHAIN 602
FT CHAIN 603
FT CHAIN 604
FT CHAIN 605
FT CHAIN 606
FT CHAIN 607
FT CHAIN 608
FT CHAIN 609
FT CHAIN 610
FT CHAIN 611
FT CHAIN 612
FT CHAIN 613
FT CHAIN 614
FT CHAIN 615
FT CHAIN 616
FT CHAIN 617
FT CHAIN 618
FT CHAIN 619
FT CHAIN 620
FT CHAIN 621
FT CHAIN 622
FT CHAIN 623
FT CHAIN 624
FT CHAIN 625
FT CHAIN 626
FT CHAIN 627
FT CHAIN 628
FT CHAIN 629
FT CHAIN 630
FT CHAIN 631
FT CHAIN 632
FT CHAIN 633
FT CHAIN 634
FT CHAIN 635
FT CHAIN 636
FT CHAIN 637
FT CHAIN 638
FT CHAIN 639
FT CHAIN 640
FT CHAIN 641
FT CHAIN 642
FT CHAIN 643
FT CHAIN 644
FT CHAIN 645
FT CHAIN 646
FT CHAIN 647
FT CHAIN 648
FT CHAIN 649
FT CHAIN 650
FT CHAIN 651
FT CHAIN 652
FT CHAIN 653
FT CHAIN 654
FT CHAIN 655
FT CHAIN 656
FT CHAIN 657
FT CHAIN 658
FT CHAIN 659
FT CHAIN 660
FT CHAIN 661
FT CHAIN 662
FT CHAIN 663
FT CHAIN 664
FT CHAIN 665
FT CHAIN 666
FT CHAIN 667
FT CHAIN 668
FT CHAIN 669
FT CHAIN 670
FT CHAIN 671
FT CHAIN 672
FT CHAIN 673
FT CHAIN 674
FT CHAIN 675
FT CHAIN 676
FT CHAIN 677
FT CHAIN 678
FT CHAIN 679
FT CHAIN 680
FT CHAIN 681
FT CHAIN 682
FT CHAIN 683
FT CHAIN 684
FT CHAIN 685
FT CHAIN 686
FT CHAIN 687
FT CHAIN 688
FT CHAIN 689
FT CHAIN 690
FT CHAIN 691
FT CHAIN 692
FT CHAIN 693
FT CHAIN 694
FT CHAIN 695
FT CHAIN 696
FT CHAIN 697
FT CHAIN 698
FT CHAIN 699
FT CHAIN 700
FT CHAIN 701
FT CHAIN 702
FT CHAIN 703
FT CHAIN 704
FT CHAIN 705
FT CHAIN 706
FT CHAIN 707
FT CHAIN 708
FT CHAIN 709
FT CHAIN 710
FT CHAIN 711
FT CHAIN 712
FT CHAIN 713
FT CHAIN 714
FT CHAIN 715
FT CHAIN 716
FT CHAIN 717
FT CHAIN 718
FT CHAIN 719
FT CHAIN 720
FT CHAIN 721
FT CHAIN 722
FT CHAIN 723
FT CHAIN 724
FT CHAIN 725
FT CHAIN 726
FT CHAIN 727
FT CHAIN 728
FT CHAIN 729
FT CHAIN 730
FT CHAIN 731
FT CHAIN 732
FT CHAIN 733
FT CHAIN 734
FT CHAIN 735
FT CHAIN 736
FT CHAIN 737
FT CHAIN 738
FT CHAIN 739
FT CHAIN 740
FT CHAIN 741
FT CHAIN 742
FT CHAIN 743
FT CHAIN 744
FT CHAIN 745
FT CHAIN 746
FT CHAIN 747
FT CHAIN 748
FT CHAIN 749
FT CHAIN 750
FT CHAIN 751
FT CHAIN 752
FT CHAIN 753
FT CHAIN 754
FT CHAIN 755
FT CHAIN 756
FT CHAIN 757
FT CHAIN 758
FT CHAIN 759
FT CHAIN 760
FT CHAIN 761
FT CHAIN 762
FT CHAIN 763
FT CHAIN 764
FT CHAIN 765
FT CHAIN 766
FT CHAIN 767
FT CHAIN 768
FT CHAIN 769
FT CHAIN 770
FT CHAIN 771
FT CHAIN 772
FT CHAIN 773
FT CHAIN 774
FT CHAIN 775
FT CHAIN 776
FT CHAIN 777
FT CHAIN 778
FT CHAIN 779
FT CHAIN 780
FT CHAIN 781
FT CHAIN 782
FT CHAIN 783
FT CHAIN 784
FT CHAIN 785
FT CHAIN 786
FT CHAIN 787
FT CHAIN 788
FT CHAIN 789
FT CHAIN 790
FT CHAIN 791
FT CHAIN 792
FT CHAIN 793
FT CHAIN 794
FT CHAIN 795
FT CHAIN 796
FT CHAIN 797
FT CHAIN 798
FT CHAIN 799
FT CHAIN 800
FT CHAIN 801
FT CHAIN 802
FT CHAIN 803
FT CHAIN 804
FT CHAIN 805
FT CHAIN 806
FT CHAIN 807
FT CHAIN 808
FT CHAIN 809
FT CHAIN 810
FT CHAIN 811
FT CHAIN 812
FT CHAIN 813
FT CHAIN 814
FT CHAIN 815
FT CHAIN 816
FT CHAIN 817
FT CHAIN 818
FT CHAIN 819
FT CHAIN 820
FT CHAIN 821
FT CHAIN 822
FT CHAIN 823
FT CHAIN 824
FT CHAIN 825
FT CHAIN 826
FT CHAIN 827
FT CHAIN 828
FT CHAIN 829
FT CHAIN 830
FT CHAIN 831
FT CHAIN 832
FT CHAIN 833
FT CHAIN 834
FT CHAIN 835
FT CHAIN 836
FT CHAIN 837
FT CHAIN 838
FT CHAIN 839
FT CHAIN 840
FT CHAIN 841
FT CHAIN 842
FT CHAIN 843
FT CHAIN 844
FT CHAIN 845
FT CHAIN 846
FT CHAIN 847
FT CHAIN 848
FT CHAIN 849
FT CHAIN 850
FT CHAIN 851
FT CHAIN 852
FT CHAIN 853
FT CHAIN 854
FT CHAIN 855
FT CHAIN 856
FT CHAIN 857
FT CHAIN 858
FT CHAIN 859
FT CHAIN 860
FT CHAIN 861
FT CHAIN 862
FT CHAIN 863
FT CHAIN 864
FT CHAIN 865
FT CHAIN 866
FT CHAIN 867
FT CHAIN 868
FT CHAIN 869
FT CHAIN 870
FT CHAIN 871
FT CHAIN 872
FT CHAIN 873
FT CHAIN 874
FT CHAIN 875
FT CHAIN 876
FT CHAIN 877
FT CHAIN 878
FT CHAIN 879
FT CHAIN 880
FT CHAIN 881
FT CHAIN 882
FT CHAIN 883
FT CHAIN 884
FT CHAIN 885
FT CHAIN 886
FT CHAIN 887
FT CHAIN 888
FT CHAIN 889
FT CHAIN 890
FT CHAIN 891
FT CHAIN 892
FT CHAIN 893
FT CHAIN 894
FT CHAIN 895
FT CHAIN 896
FT CHAIN 897
FT CHAIN 898
FT CHAIN 899
FT CHAIN 900
FT CHAIN 901
FT CHAIN 902
FT CHAIN 903
FT CHAIN 904
FT CHAIN 905
FT CHAIN 906
FT CHAIN 907
FT CHAIN 908
FT CHAIN 909
FT CHAIN 910
FT CHAIN 911
FT CHAIN 912
FT CHAIN 913
FT CHAIN 914
FT CHAIN 915
FT CHAIN 916
FT CHAIN 917
FT CHAIN 918
FT CHAIN 919
FT CHAIN 920
FT CHAIN 921
FT CHAIN 922
FT CHAIN 923
FT CHAIN 924
FT CHAIN 925
FT CHAIN 926
FT CHAIN 927
FT CHAIN 928
FT CHAIN 929
FT CHAIN 930
FT CHAIN 931
FT CHAIN 932
FT CHAIN 933
FT CHAIN 934
FT CHAIN 935
FT CHAIN 936
FT CHAIN 937
FT CHAIN 938
FT CHAIN 939
FT CHAIN 940
FT CHAIN 941
FT CHAIN 942
FT CHAIN 943
FT CHAIN 944
FT CHAIN 945
FT CHAIN 946
FT CHAIN 947
FT CHAIN 948
FT CHAIN 949
FT CHAIN 950
FT CHAIN 951
FT CHAIN 952
FT CHAIN 953
FT CHAIN 954
FT CHAIN 955
FT CHAIN 956
FT CHAIN 957
FT CHAIN 958
FT CHAIN 959
FT CHAIN 960
FT CHAIN 961
FT CHAIN 962
FT CHAIN 963
FT CHAIN 964
FT CHAIN 965
FT CHAIN 966
FT CHAIN 967
FT CHAIN 968
FT CHAIN 969
FT CHAIN 970
FT CHAIN 971
FT CHAIN 972
FT CHAIN 973
FT CHAIN 974
FT CHAIN 975
FT CHAIN 976
FT CHAIN 977
FT CHAIN 978
FT CHAIN 979
FT CHAIN 980
FT CHAIN 981
FT CHAIN 982
FT CHAIN 983
FT CHAIN 984
FT CHAIN 985
FT CHAIN 986
FT CHAIN 987
FT CHAIN 988
FT CHAIN 989
FT CHAIN 990
FT CHAIN 991
FT CHAIN 992
FT CHAIN 993
FT CHAIN 994
FT CHAIN 995
FT CHAIN 996
FT CHAIN 997
FT CHAIN 998
FT CHAIN 999
FT CHAIN 1000
FT CHAIN 1001
FT CHAIN 1002
FT CHAIN 1003
FT CHAIN 1004
FT CHAIN 1005
FT CHAIN 1006
FT CHAIN 1007
FT CHAIN 1008
FT CHAIN 1009
FT CHAIN 1010
FT CHAIN 1011
FT CHAIN 1012
FT CHAIN 1013
FT CHAIN 1014
FT CHAIN 1015
FT CHAIN 1016
FT CHAIN 1017
FT CHAIN 1018
FT CHAIN 1019
FT CHAIN 1020
FT CHAIN 1021
FT CHAIN 1022
FT CHAIN 1023
FT CHAIN 1024
FT CHAIN 1025
FT CHAIN 1026
FT CHAIN 1027
FT CHAIN 1028
FT CHAIN 1029
FT CHAIN 1030
FT CHAIN 1031
FT CHAIN 1032
FT CHAIN 1033
FT CHAIN 1034
FT CHAIN 1035
FT CHAIN 1036
FT CHAIN 1037
FT CHAIN 1038
FT CHAIN 1039
FT CHAIN 1040
FT CHAIN 1041
FT CHAIN 1042

```

DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).  
OS Hepatitis A virus (strain CR326).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.  
ON NCBI\_TaxID=12097;  
RX MEDLINE=91162758; PubMed=2985793;  
RA Linemeyer D.L., Wenke J.G., Martin-Gallardo A., Hughes J.V., Young A., Mitra S.W.;  
RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA";  
RL J. Virol. 54:247-255(1985).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS.  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.  
CC -1- SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M10033; AAA45470.1; -.  
DR PIR: A03904; GNNYHA.  
KW Polypeptide; Coat protein; Core protein.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 836  
FT CHAIN 837 >852  
FT NON\_TER 852 852  
SQ SEQUENCE 852 AA; 95563 MW; 73D3BD0AD532820E CRC64;  
Query Match 100.0%; Score 107; DB 1; Length 852;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SHIECRKPYKELRLGVGKOR 20  
Db 792 SHIECRKPYKELRLGVGKOR 811  
RESULT 3  
ID POLG\_HPAV2 STANDARD; PRT; 2226 AA.  
AC P26580;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
DE P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain 24a).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.  
ON NCBI\_TaxID=12094;  
RX MEDLINE=91162758; PubMed=1705995;  
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M., Crommons T., Jansen R.W.;  
RT "Antigenic and genetic variation in cytopathic hepatitis A virus RT variants arising during persistent infection: evidence for genetic recombination";  
RL J. Virol. 65:2056-2065(1991).

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate + RNA(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS.  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.  
CC -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M59810; AAA45468.1; -.  
DR MEROPS: C03.005; -.  
DR InterPro: IPR004004; Calic\_pol\_hel.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR007094; RNA\_pol\_PSV1r.  
DR Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam: PF00910; RNA\_helicase; 1.  
DR PRINTS: PR00918; CALICVIRUSNS.  
KW Polypeptide; Coat protein; Core protein; Transferrase; RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 794  
FT CHAIN 795 900  
FT CHAIN 901 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1495  
FT CHAIN 1496 1518  
FT CHAIN 1519 1737  
FT CHAIN 1738 2226  
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6BA4E2BF CRC64;  
Query Match 100.0%; Score 107; DB 1; Length 2226;  
Best Local Similarity 100.0%; Pred. No. 4.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 SHIECRKPYKELRLGVGKOR 20  
Db 792 SHIECRKPYKELRLGVGKOR 811  
RESULT 4  
ID POLG\_HPAV4 STANDARD; PRT; 2226 AA.  
AC P26581;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
DE P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain 43c).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.  
ON NCBI\_TaxID=12095;  
RX MEDLINE=91162758; PubMed=1705995;  
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M., Crommons T., Jansen R.W.;  
RT "Antigenic and genetic variation in cytopathic hepatitis A virus RT variants arising during persistent infection: evidence for genetic recombination";  
RL J. Virol. 65:2056-2065(1991).

```
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC (RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: M59809; AAA45469.1; -.  
DR MEROPS: C03.005; -.  
DR InterPro: IPR000605; RNA_helicase.  
DR InterPro: IPR007095; RNA_pol_DS_PS.  
DR InterPro: IPR001205; RNA_pol_P3D.  
DR InterPro: IPR007094; RNA_pol_PSVlr.  
DR Pfam: PF00680; RNA_dep_RNA_pol: 1.  
DR Pfam: PF00910; RNA_helicase: 1.  
DR Polyprotein; Coat protein; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease;  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 794  
FT CHAIN 795 900  
FT CHAIN 901 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1495  
FT CHAIN 1496 1518  
FT CHAIN 1519 1737  
FT CHAIN 1738 2226  
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80509BF75 CRC64;  
  
Query Match 100.0%; Score 107; DB 1; Length 2226;  
Best Local Similarity 100.0%; Pred. No. 4.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 SHIECRKPYKELRLVGVKOR 20  
DB 792 SHIECRKPYKELRLVGVKOR 811  
  
RESULT 5  
POLG_HPAV8 STANDARD; PRT; 2226 AA.  
ID POLG_HPAV8  
AC P26582;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins  
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
DE P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain 18f).  
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI_TaxID=12096;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91162758; PubMed=1705995;  
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstein S.M.,  
RA Cromeans T., Jansen R.W.;  
RT "Antigenic and genetic variation in cytopathic hepatitis A virus  
RT variants arising during persistent infection: evidence for genetic  
RT recombination.";  
RL J. Virol. 65:2056-2065(1991).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC (RNA)(N).
```

```
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: M59808; AAA45467.1; -.  
DR PDB: 1OAT; 15-MAY-00.  
DR MEROPS: C03.005; -.  
DR InterPro: IPR000605; RNA_helicase.  
DR InterPro: IPR007095; RNA_pol_DS_PS.  
DR InterPro: IPR001205; RNA_pol_P3D.  
DR InterPro: IPR007094; RNA_pol_PSVlr.  
DR Pfam: PF00680; RNA_dep_RNA_pol: 1.  
DR Pfam: PF00910; RNA_helicase: 1.  
DR Polyprotein; Coat protein; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure;  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 794  
FT CHAIN 795 900  
FT CHAIN 901 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1495  
FT CHAIN 1496 1518  
FT CHAIN 1519 1737  
FT CHAIN 1738 2226  
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;  
  
Query Match 100.0%; Score 107; DB 1; Length 2226;  
Best Local Similarity 100.0%; Pred. No. 4.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 SHIECRKPYKELRLVGVKOR 20  
DB 792 SHIECRKPYKELRLVGVKOR 811  
  
RESULT 6  
POLG_HPAV8 STANDARD; PRT; 2227 AA.  
ID POLG_HPAV8  
AC P08617; P06443; O81082;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins  
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
DE P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain HM-175).  
OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI_TaxID=12098;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87061253; PubMed=3023706;  
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,  
RA Baroudy B.M.;  
RT "Complete nucleotide sequence of wild-type hepatitis A virus:  
RT comparison with different strains of hepatitis A virus and other  
RT picornaviruses.";  
RL J. Virol. 61:50-59(1987).  
RN [2]  
RP SEQUENCE FROM N.A.
```

RC STRAIN-Attenuated;  
RX MEDLINE=87175701; PubMed=3031686;  
RA Cohen J.T., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M.,  
RT Purcell R.H.;  
RT "Complete nucleotide sequence of an attenuated hepatitis A virus:  
RL comparison with wild-type virus.";  
RN Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).  
RP [3]  
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.  
RX MEDLINE=85166289; PubMed=2984684;  
RA Barouly B.M., Ticehurst J.R., Mele T.A., Maize J.V. Jr.,  
RT Purcell R.H., Feinstone S.M.;  
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid  
RL proteins and RNA polymerase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).  
CC -I- CAPSALTY ACTIVITY: N nucleoside triphosphate = N diphosphate. +  
CC (RNA)(N).  
CC -I- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -I- PRIM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATRE PROTEINS.  
CC -I- MISCELLANEOUS: STRAIN HM-175/77 MK-5 IS ATTENUATED STRAIN DERIVED  
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.  
CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -I- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT  
CC SHOWN.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on ways  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
CC -----  
CC EMBL; M14114; AAA45475.1; -;  
DR EMBL; M14707; AAA45465.1; -;  
DR EMBL; M14707; AAA45466.1; ALT\_INT.  
DR EMBL; M16632; AAA45471.1; -;  
DR PIR; A25981; GNNYTM.  
DR PIR; A94149; GNNYTM.  
DR PDB; 1HAV; 23-DEC-96.  
DR MEROPS; C03.005; -;  
DR InterPro: IPR004004; Calic.pol.bel.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR007094; RNA\_pol\_PSVr.  
Pfam: PF00680; RNA\_dep\_RNA\_pol. 1.  
Pfam: PF00910; RNA\_helicase; 1.  
PRINTS; PR00918; CALICYRUSNS.  
KW Polyprotein; Coat protein; Core protein; Transferase;  
RM RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 836  
FT CHAIN 837 980  
FT CHAIN 981 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1496  
FT CHAIN 1497 1519  
FT CHAIN 1520 1738  
FT CHAIN 1739 2227  
FT CHAIN 77 77  
FT VARIANT 764 764  
FT VARIANT 821 821  
FT VARIANT 1052 1052  
FT VARIANT 1062 1062  
FT VARIANT 1118 1118  
FT VARIANT 1151 1151  
FT VARIANT 1163 1163  
FT VARIANT 1277 1277  
FT V -> I (IN ATTENUATED STRAIN).  
E -> V (IN ATTENUATED STRAIN).  
N -> S (IN ATTENUATED STRAIN).  
A -> V (IN ATTENUATED STRAIN).  
G -> A (IN ATTENUATED STRAIN).  
K -> M (IN ATTENUATED STRAIN).  
E -> K (IN ATTENUATED STRAIN).  
F -> S (IN ATTENUATED STRAIN).  
V -> I (IN ATTENUATED STRAIN).

[illegible]

FT CHAIN 1508 1678 PROBABLE PROTEIN P3C.  
 FT CHAIN 1679 2227 RNA-DIRECTED POLYMERASE P3D.  
 SQ SEQUENCE 2227 AA: 251898 MW: 99A7354B4CD2799C CRC64:

Query Match 100.0%; Score 107; DB 1; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECKRPPKELRLLEVGR 20  
 DB 792 SHIECKRPPKELRLLEVGR 811

RESULT 8  
 ID POLG\_HPAVM STANDARD; PRT: 2227 AA.  
 AC P13901: 081083; 081084; 081085; 081086; 081087; 081088; 081089;  
 AC 081090; 081091; 081092; 081093;  
 DT 01-JAN-1990 (rel. 13, Created)  
 DT 01-JAN-1990 (rel. 13, Last sequence update)  
 DT 28-FEB-2003 (rel. 41, Last annotation update)  
 DE Genome polypeptide [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
 DE Hepatitis A virus (strain MBB).  
 OS Hepatitis A virus (strain MBB).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Hepatovirus.  
 OX NCBI\_TaxID=12100;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88045071; PubMed=2823500;  
 RA Paul A.V., Tada H., der Helm K., Wissel T., Klein R., Wimmer E.,  
 RA Delnhardt F.;  
 RT The entire nucleotide sequence of the genome of human hepatitis A virus (isolate MBB).  
 RT Virus Res 8:153-171(1987).  
 RL -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + (RNA)(N).  
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.  
 CC -1- PTM: SPECIFIC ENZYMTIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).  
 CC -----  
 CC EMBL: M20273; AAA45474.1; -  
 DR MEROPS: C03.005; -  
 DR InterPro: IPR000605; RNA\_helicase.  
 DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro: IPR001205; RNA\_pol\_P3D.  
 DR InterPro: IPR007094; RNA\_pol\_PSVIT.  
 DR Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR PolyProtein: Coat protein; Core protein; Transferase;  
 DR RNA-directed RNA polymerase; Hydrolyase; Thiol protease.  
 KW CHAIN 1 23  
 FT CHAIN 24 245 COAT PROTEIN VP4 (P1A).  
 FT CHAIN 24 491 COAT PROTEIN VP2 (P1B).  
 FT CHAIN 492 836 COAT PROTEIN VP3 (P1C).  
 FT CHAIN 837 980 COAT PROTEIN VP1 (P1D).  
 FT CHAIN 981 1087 CORE PROTEIN P2A.  
 FT CHAIN 1088 1422 CORE PROTEIN P2B.  
 FT CHAIN 1423 1496 CORE PROTEIN P2C.  
 FT CHAIN 1497 1519 PROBABLE PROTEIN P3A.  
 FT CHAIN 1519 PROBABLE PROTEIN P3B.

FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.  
 FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.  
 SQ SEQUENCE 2227 AA: 251425 MW: EC983ED2A7C86349 CRC64:

Query Match 100.0%; Score 107; DB 1; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 4.9e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECKRPPKELRLLEVGR 20  
 DB 792 SHIECKRPPKELRLLEVGR 811

RESULT 9  
 ID POLG\_HPAVM STANDARD; PRT: 808 AA.  
 AC 002381;  
 DT 01-JUL-1993 (rel. 26, Created)  
 DT 01-JUL-1993 (rel. 26, Last sequence update)  
 DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE Genome polypeptide [contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).  
 DE Hepatitis A virus (strain GA76).  
 OS Hepatitis A virus (strain GA76).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Hepatovirus.  
 OX NCBI\_TaxID=31706;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92260183; PubMed=1316423;  
 RA Khanna B., Spelbring J.E., Innis B.L., Robertson B.H.;  
 RT Characterization of a genetic variant of human hepatitis A virus.\*;  
 RL J. Med. Virol. 36:118-124(1992).  
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS. EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.  
 CC -1- PTM: SPECIFIC ENZYMTIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).  
 CC -----  
 CC EMBL: M6695; AAA45477.1; -  
 DR PolyProtein: Coat protein; Core protein.  
 KW CHAIN 1 1  
 FT CHAIN <1 2 COAT PROTEIN VP4 (P1A).  
 FT CHAIN 3 223 COAT PROTEIN VP2 (P1B).  
 FT CHAIN 224 470 COAT PROTEIN VP3 (P1C).  
 FT CHAIN 471 770 COAT PROTEIN VP1 (P1D).  
 FT CHAIN 771 >808 CORE PROTEIN P2A.  
 FT NON\_TER 808 808  
 SQ SEQUENCE 808 AA: 90632 MW: D80CE7E57A479C12 CRC64:

Query Match 88.8%; Score 95; DB 1; Length 808;  
 Best Local Similarity 95.0%; Pred. No. 1.5e-07;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 SHIECKRPPKELRLLEVGR 20  
 DB 771 SHIECKRPPKELRLLEVGR 790

RESULT 10  
 ID POLG\_HPAVM STANDARD; PRT: 2230 AA.  
 AC P14553;  
 DT 01-JAN-1990 (rel. 13, Created)  
 DT 01-AUG-1992 (rel. 23, Last sequence update)  
 DT 28-FEB-2003 (rel. 41, Last annotation update)  
 DE Genome polypeptide [contains: Coat proteins VP1 TO VP4; Core proteins

DE	P2A TO P2C; probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE	P3D (EC 2.7.7.48)].
OS	Simian hepatitis A virus (strain AGM-27).
OC	Viruses; ssRNA positive strand viruses, no DNA stage; Picornaviridae;
CC	Hepadnavirus.
ON	NCB1_TaxID=12102;
RX	SEQUENCE FROM N.A.
RX	MEDLINE=913111420; PubMed=1649901;
RA	Tsarev S.A., Emerson S.U., Balayan M.S., Tichehurst J.R.,
RA	Purell R.H.;
RT	"Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RL	structure and growth in cell culture with other HAV strains.";
RN	J. Gen. Virol. 72:1677-1683(1991).
[2]	
RP	SEQUENCE OF 1750-2164 FROM N.A.
RX	MEDLINE=89232168; PubMed=2541023;
RA	Balayan M.S., Kusov Y.Y., Andapaidze A.G., Tsarev S.A.;
RA	Sverdlov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
CC	"Variations in genome fragments coding for RNA polymerase in human
CC	and simian hepatitis A viruses.";
CC	FEBS Lett. 247:425-428(1989).
CC	-1- CATALYTIC ACTIVITY: N nucleoside triphosphate +
CC	(RNA).(N).
CC	-1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC	EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC	VP3, AND VP4.
CC	-1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC	-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see <a href="http://www.isb-sdb.ch/announce/">http://www.isb-sdb.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sdb.ch">license@isb-sdb.ch</a> ).
CC	-----
DR	EMBL, D00924; BAA00766.1; -
DR	EMBL, X15461; CA33490.1; -
DR	PIR, A30470; GNNYSA.
DR	MEROPS: C03.005; -
DR	InterPro: IPR004004; Calic1_pol_hel.
DR	InterPro: IPR000605; RNA_helicase.
DR	InterPro: IPR007095; RNA_pol_DS_Pf.
DR	InterPro: IPR001205; RNA_pol_P3D.
DR	InterPro: IPR007094; RNA_pol_psv1r.
DR	Pfam: PF00680; RNA_dep_RNA_pol_1.
DR	Pfam: PF00910; RNA_helicase_1.
DR	PRINTS: PR00918; CALICVIRDSNS.
KM	Polyprotein; Coat protein; Core protein; Hydroxase; Thiol protease.
FT	CHAIN 1 27
FT	CHAIN 28 249
FT	CHAIN 250 495
FT	CHAIN 496 795
FT	CHAIN 796 984
FT	CHAIN 985 1091
FT	CHAIN 1092 1426
FT	CHAIN 1427 1498
FT	CHAIN 1499 1521
FT	CHAIN 1522 1741
FT	CHAIN 1742 2230
SQ	SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;
QY	Query Match 80.4%; Score 86; DB 1; Length 2230;
.	Best Local Similarity 85.0%; Pred. No. 1.2e+05;
Matches	17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
DB	1 SHIECRKPYKELRLEVGROR 20
	796 SHIEGKRYKELRREVGROR 815

```

RESULT 11
POLG_HPAVT STANDARD: PRT: 839 AA.
AC P31788;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core protein
P2a] (Fragment).
OS Simian hepatitis A virus (strain CY-145).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=31707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91311421; PubMed=1649902;
RA Nathan O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;
RT "Sequence analysis of a new hepatitis A virus naturally infecting
RT cynomolgus macaques (Macaca fascicularis).";
RL J. Gen. Virol. 72:1685-1689(1991)
-1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
-1- PWM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb.ch/announce/
CC CC or send an email to license@isb.ch).
-----
DR EMBL; M59286; AAA5473.1; -
DR PIR: J01180; GNNS2.
KW Polypeptide; Coat protein; Core protein.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 ? COAT PROTEIN VP1 (PID).
FT CHAIN ? >839 CORE PROTEIN P2A.
FT NON_TER 839 839
SQ SEQUENCE 839 AA; 93825 MW; 2CACCA4BD1E192DBC CRC64;
Query Match 79.0%; Score 84.5; DB 1; Length 839;
Best Local Similarity 90.0%; Pred. No. 7.7e-06;
Matches 18; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 SHIECRKPKYKELREVGKOR 20
Db 792 SHIE-KPKYKELREVGKOR 810
IIII:IIIIIIIIIIII
RESULT 12
YAGA_SCHPO STANDARD: PRT: 420 AA.
AC Q09873;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C12G12.10 in chromosome I.
GN SPAC12G12.10.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,

```

RA Squire J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Holtroyd S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holtroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA Jones K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Moore P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tiley A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkart G., Aert R., Robben J., Grymoprez B.,  
 RA Wajters I., Vansireels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Medler H., Mambutt R., Punelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaue V., Motier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerruti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,  
 RA Shpakovskii G.V., Ussery D., Barrett B.G., Nurse P.,  
 RA "The genome sequence of Schizosaccharomyces pombe".  
 RL Nature 415:871-880(2002).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).

CC EMBL: Z66568; CAA91505.1; -  
 DR PIR: S62541; S62541.  
 DR GenBank: SPAC12G12.10; -  
 DR InterPro: IPR001680; WD40.  
 DR Pfam: PF00400; WD40; 1.  
 DR SMART: SM00320; WD40; 3.  
 DR Hypothetical protein.  
 KW SEQUENCE 420 AA; 47525 MW; 9DC1F710FBEB9761 CRC64;

Query Match 43.0%; Score 46; DB 1; Length 420;  
 Best Local Similarity 46.7%; Pred. No. 6.2;  
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

4 ECRPKYKELREYVGK 18  
 :||:|:|:|:|:|:|  
 376 DCSLPFKEMRYDDGK 390

RESULT 13  
 GCP6\_HUMAN STANDARD; PRT; 1819 AA.  
 ID GCP6\_HUMAN STANDARD; PRT; 1819 AA.  
 AC Q96RT7; Q9BY91; Q9UGX3; Q9UGX4;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Gamma-tubulin complex component 6 (GCP-6).  
 GN TUBGCP6 OR GCP6 OR KIAA1659.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1) AND CHARACTERIZATION.  
 RX MEDLINE=21551508; PubMed=11694571.  
 RA Murphy S.M., Preble A.M., Patel U.K., Dias D.P.,  
 RA Murphy M., Agard D., Stuifjes J.T., Stearns T.,  
 RT "GCP5 and GCP6: two new members of the human gamma-tubulin complex".  
 RL Mol. Biol. Cell 12:3340-3352(2001).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=20057165; PubMed=10591208;  
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,  
 RA Clump M., Smink L.J., Atncough R., Almeida J.P., Babbage A.K.,  
 RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasly O.P.,  
 RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,  
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,  
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,  
 RA Conroy D., Cockby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,  
 RA Dhami P.D., Dockrey C., Dodsworth S.J., Durbin R.M., Ellington A.G.,  
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,  
 RA Gilbert J.G.R., Goward M.E., Grafton D.V., Griffiths M.N.D., Hall C.,  
 RA Hall R.E., Hall-Famlyn G., Heathcote R.W., Ho S., Holmes S.,  
 RA Laird S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,  
 RA Laird G.K., Langford C.F., Leverisha M.A., Lloyd C., Lloyd D.M.,  
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,  
 RA Mcclay J., McLaren S., McMurtry A.A., Milne S.A., Mortimore B.J.,  
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,  
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,  
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,  
 RA Steward C.A., Steward C.A., Sulston J.E., Swann R.M.,  
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,  
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,  
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,  
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,  
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,  
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,  
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malay E., Nguyen T., Pan H.,  
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,  
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,  
 RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,  
 RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,  
 RA Corcos M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,  
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,  
 RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,  
 RA Kort I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,  
 RA Emden H.S., Shaikh T., Kurahashi H., Saita S., Budarf M.L.,  
 RA Kim U.J., Shizuya H., Simon M.I., Dumanowski J.P., Peyrard M., Kedra D.,  
 RA Seroussi E., Franssion I., Tapia I., Brudner C.F., O'Brien K.P.,  
 RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,  
 RA Tiliun Y., Wright H.,  
 RT "The DNA sequence of human chromosome 22".  
 RL Nature 402:489-495(1999).  
 RN [3]  
 RP SEQUENCE OF 33-1819 FROM N.A. (ISOFORM 2).  
 RC TISSUE=Brain;  
 RX MEDLINE=21156230; PubMed=11258795;  
 RA Hirotsawa M., Nagase T., Murahashi Y., Kikuno R., Ohara O.,  
 RT "Identification of novel transcribed sequences on human chromosome 22  
 RT by expressed sequence tag mapping".  
 RL DNA Res. 8:1-9(2001).  
 CC -1- FUNCTION: Gamma-tubulin complex is necessary for microtubule  
 CC nucleation at the centrosome.  
 CC -1- SUBUNIT: Gamma-tubulin complex is composed of gamma-tubulin, GCP2,  
 CC GCP3, GCP4, GCP5 and GCP6.  
 CC -1- SUBCELLULAR LOCATION: Centrosome.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=Q96RT7-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=Q96RT7-2; Sequence=VSP\_001624;  
 CC Note=No experimental confirmation available;  
 CC -1- SIMILARITY: BELONGS TO THE GCP FAMILY.  
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to frameshifts  
 CC in positions 1371 and 1758.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF272887; AAK82968.1; -
DR EMBL; AL022328; CAB63046.1; ALT_SEQ.
DR EMBL; AL022328; CAB63047.1; ALT_SEQ.
DR EMBL; AB051456; BAB33339.1; ALT_FRAME.
DR Genev; HGNC:18127; TUBSCP6.
DR GO; GO:0008274; C:gamma-tubulin ring complex; IDA.
DR GO; GO:0008017; F:microtubule binding activity; IDA.
DR GO; GO:0007020; P:microtubule nucleation; IDA.
DR Pfam; PF04130; Spc97_Spc98; 1.
KW Microtubules; Repeat; Alternative splicing.
FT DOMAIN 1027 1269 9 x 27 AA TANDEM REPEATS.
FT REPEAT 1027 1053 1.
FT REPEAT 1054 1080 2.
FT REPEAT 1081 1107 3.
FT REPEAT 1108 1134 4.
FT REPEAT 1135 1161 5.
FT REPEAT 1162 1188 6.
FT REPEAT 1189 1215 7.
FT REPEAT 1216 1242 8.
FT REPEAT 1243 1269 9.
FT VARSPLIC 1724 1757 Missing (In Isoform 2).
FT CONFLICT 567 567 /FTID-VSP_001624.
FT CONFLICT 1377 1377 S -> L (IN REF. 3).
FT CONFLICT 1621 1621 A -> T (IN REF. 2 AND 3).
SQ SEQUENCE 1819 AA; 200455 MW; 923576544D3454A CRC64;
Query Match Similarity 42.1%; Score 45; DB 1; Length 1819;
Best Local 50.0%; Pred. NO. 39;
Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
QY 2 HIECRKPYKELRELVGKQ 19
Db 653 HSSVSKKEKELMEIAKQ 670
-----
RESULT 14
CO3_RAT
ID CO3_RAT STANDARD; PRT; 1663 AA.
AC P01026;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Complement C3 precursor [contains: C3a anaphylatoxin].
C3
Rattus norvegicus (Rat).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=Wistar; TISSUE=Liver;
RC MEDLINE=90245672; PubMed=2336397;
RA Misumi Y., Sohma M., Ikehara Y.;
RT "Nucleotide and deduced amino acid sequence of rat complement C3.",
RL Nucleic Acids Res. 18:2178-2178(1990).
RN [2]
RP SEQUENCE OF 671-748.
RX MEDLINE=79062262; PubMed=309768;
RA Jacobs J.W., Rubin J.S., Hugli T.E., Bogardt R.A., Mariz I.K.,
RA Daniels J.S., Raghunaday W.H., Birdsaw R.A.;
RT "Purification, characterization, and amino acid sequence of rat
anaphylatoxin (C3a).";
RL Biochemistry 17:5031-5038(1978).
RN [3]
RP SEQUENCE OF 1316-1595 FROM N.A.
RX MEDLINE=89380332; PubMed=2674144;
RA Sundstrom S.A., Konni B.S., Ponce-De-Leon H., Yi Z., Teuscher C.,
RA Lytle C.R.;

```

RT		"Steroidregulation of tissue-specific expression of complement C3.";
RL	J. Biol. Chem. 264:16941-16947(1989)	
CC	-1- FUNCTION: C3 PLAYS A CENTRAL ROLE IN THE ACTIVATION OF THE	
CC	COMPLEMENT SYSTEM. THIS PROCESSING BY C3 CONVERTASE IS THE CENTRAL	
CC	REACTION IN BOTH CLASSICAL AND ALTERNATIVE COMPLEMENT PATHWAYS.	
CC	AFTER ACTIVATION C3B CAN BIND COVALENTLY, VIA ITS REACTIVE	
CC	THIOESTER, TO CELL SURFACE CARBOHYDRATES OR IMMUNE AGGREGATES.	
CC	-1- FUNCTION: DERIVED FROM PROTEOLYTIC DEGRADATION OF COMPLEMENT C3,	
CC	C3A ANAPHYLATOXIN IS A MEDIATOR OF LOCAL INFLAMMATORY PROCESS. IT	
CC	INDUCES THE CONTRACTION OF SMOOTH MUSCLE, INCREASES VASCULAR	
CC	PERMEABILITY AND CAUSES HISTAMINE RELEASE FROM MAST CELLS AND	
CC	BASOPHILIC LEUCOCYTES.	
CC	-1- SUBUNIT: C3 PRECURSOR IS FIRST PROCESSED BY THE REMOVAL OF 4 ARG	
CC	RESIDUES FORMING TWO CHAINS, BETA & ALPHA, LINKED BY A DISULFIDE	
CC	BOND. C3 CONVERTASE ACTIVATES C3 BY CLEAVING THE ALPHA CHAIN,	
CC	RELEASENING C3A ANAPHYLATOXIN & GENERATING C3B (BETA CHAIN + ALPHA	
CC	CHAIN).	
CC	-1- SIMILARITY: TO C4, C5 AND ALPHA-2-MACROGLOBULIN.	
CC	-1- SIMILARITY: Contains 1 anaphylatoxin-like domain.	
CC		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC		
DR	EMBL; X52477; CAA36716.1; -	
DR	EMBL; M29866; AAA40837.1; ALT_SEQ.	
DR	PIR; S15764; C3RP.	
DR	PDB; 1QOF; 31-JUL-00.	
DR	PDB; 1QSF; 31-JUL-00.	
DR	InterPro: IPR002890; A2M_N.	
DR	InterPro: IPR000020; Anaphylatoxin.	
DR	InterPro: IPR001599; Macrogloblna2.	
DR	InterPro: IPR001134; Netrin_C.	
DR	Pfam; PF00207; A2M; 1.	
DR	Pfam; PF01835; A2M_N; 1.	
DR	Pfam; PF01821; ANATO; 1.	
DR	Pfam; PF01759; NTR; 1.	
DR	ProDom: PD003264; Anaphylatoxin; 1.	
DR	SMART; SM00104; ANATO; 1.	
DR	SMART; SM00643; C345C; 1.	
DR	PROSITE; PS00477; ALPHA_2-MACROGLOBULIN; 1.	
DR	PROSITE; PS01177; ANAPHYLATOXIN_1; 1.	
DR	PROSITE; PS01178; ANAPHYLATOXIN_2; 1.	
KW	Complement pathway; Complement alternate pathway; Plasma;	
KW	Inflammatory response; Glycoprotein; Signal; 3D-structure;	
KW	thioester bond.	
FT	SIGNAL	24
FT	CHAIN	25
FT	CHAIN	25
FT	CHAIN	671
FT	PEPTIDE	671
FT	CHAIN	749
FT	SITE	748
FT	DOMAIN	693
FT	DISULFID	558
FT	DISULFID	626
FT	DISULFID	693
FT	DISULFID	694
FT	DISULFID	707
FT	DISULFID	873
FT	DISULFID	1101
FT	DISULFID	1158
FT	DISULFID	1358
FT	DISULFID	1389
FT	DISULFID	1506
FT	DISULFID	1518
FT	DISULFID	1537
FT	CROSSLINK	1010
FT	CARBOHYD	939
FT	CARBOHYD	1617
FT	COMPLEMENT C3.	
FT	BETA CHAIN.	
FT	ALPHA CHAIN.	
FT	C3A ANAPHYLATOXIN.	
FT	C3B (ALPHA' CHAIN).	
FT	CLEAVAGE (BY C3 CONVERTASE).	
FT	ANAPHYLATOXIN-LIKE.	
FT	INTERCHAIN (BY SIMILARITY).	
FT	BY SIMILARITY.	
FT	BY SIMILARITY.	
FT	BY SIMILARITY.	
FT	BY SIMILARITY.	
FT	BY SIMILARITY.	
FT	BY SIMILARITY.	
FT	BY SIMILARITY.	
FT	BY SIMILARITY.	
FT	Isoglutamyl cysteine thioester (Cys-Gln).	
FT	N-LINKED (GLCNAC. . . ) (PROBABLE).	
FT	N-LINKED (GLCNAC. . . ) (PROBABLE).	



FT CONFLICT 721 722 LK -> KL (IN REF. 2).  
 SQ SEQUENCE 1663 AA; 186460 MW; 2F87C8143CDDABC CRC64;

Query Match 40.7%; Score 43.5; DB 1; Length 1663;  
 Best Local Similarity 47.4%; Pred. NO. 63;

Matches 9; Conservative 4; Mismatches 3; Indels 3; Gaps 1;

OY 1 SHIECRKPYKELRLEVGRQ 19  
 ||::|| |::|||  
 DB 1586 SHVACR---NALKIQKQKQ 1601

## RESULT 15

NCAP\_HRSV1 STANDARD; PRT; 391 AA.

AC P24566;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 01-MAR-1992 (Rel. 21, Last annotation update)  
 Nucleocapsid protein.

N.

Human respiratory syncytial virus (subgroup B / strain 18537).

OC Viruses; ssRNA negative-strand viruses; Mononegavirales;

OC Paramyxoviridae; Pneumovirinae; Pneumovirus.

OX NCBI\_TaxID=11251;

[1]

SEQUENCE FROM N. A.

RX MEDLINE=89279331; PubMed=2525176;

RA Johnson P. R., Collins P. L.;

"The 1B (NS2), 1C (NS1) and N proteins of human respiratory syncytial

virus (RSV) of antigenic subgroups A and B: sequence conservation and

divergence within RSV genomic RNA.";

J. Gen. Virol. 70:1539-1547(1989).

-I- FUNCTION: MOST ABUNDANT PROTEIN IN THE VIRION AND AN IMPORTANT

ELEMENT CONFERRING HELICAL SYMMETRY ON THE NUCLEOPROTEIN CORE AS

WELL AS INTERACTING WITH M PROTEIN DURING VIRION FORMATION.

-I- SIMILARITY: BELONGS TO THE PARAMYXOVIRUSES NUCLEOCAPSID FAMILY.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

-----

DR EMBL; D00736; BAA00637.1; -

PIR; G32063; VHN23.

InterPro; IPR004930; Pneumo\_ncap.

Pfam; PF03246; Pneumo\_ncap.1.

ProDom; PD006438; Pneumo\_ncap.1.

DR Nucleocapsid

KW Nucleocapsid

SEQUENCE 391 AA; 43416 MW; 9A2BBB50103B2835 CRC64;

Query Match 40.2%; Score 43; DB 1; Length 391;

Best Local Similarity 56.2%; Pred. NO. 17;

Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 3 IECKRPYKELRLEVGR 18  
 |||||::|

DB 129 IESRSYKRLKEMGE 144

Search completed: October 1, 2003, 09:57:45  
 Job time : 9.58824 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 33.1765 seconds  
(without alignments)  
155.564 Million cell updates/sec

Title: US-09-171-432A-39

Perfect score: 107  
Sequence: 1 SHIECRKPKYKELRLEVGKOR 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Minimum number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_23:\*\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	53	12	098VY2 hepatitis a
2	107	100.0	53	12	098VY3 hepatitis a
3	107	100.0	53	12	098VY4 hepatitis a
4	107	100.0	53	12	098VY0 hepatitis a
5	107	100.0	53	12	098VY5 hepatitis a
6	107	100.0	53	12	098VY6 hepatitis a
7	107	100.0	53	12	098VY1 hepatitis a
8	107	100.0	53	12	098VY7 hepatitis a
9	107	100.0	55	12	098VY2 hepatitis a
10	107	100.0	55	12	098VY4 hepatitis a
11	107	100.0	55	12	098VY3 hepatitis a
12	107	100.0	55	12	098VY0 hepatitis a
13	107	100.0	55	12	098VY5 hepatitis a
14	107	100.0	56	12	098VY6 hepatitis a
15	107	100.0	56	12	098VY1 hepatitis a
16	107	100.0	56	12	098VY7 hepatitis a

17	107	100.0	56	12	098VY8 hepatitis a
18	107	100.0	56	12	08JYP2 hepatitis a
19	107	100.0	56	12	091PB7 hepatitis a
20	107	100.0	56	12	08JYP1 hepatitis a
21	107	100.0	56	12	098VY1 hepatitis a
22	107	100.0	56	12	098VY2 hepatitis a
23	107	100.0	56	12	08JYK3 hepatitis a
24	107	100.0	56	12	091BQ3 hepatitis a
25	107	100.0	56	12	091PB3 hepatitis a
26	107	100.0	56	12	091PB2 hepatitis a
27	107	100.0	56	12	098VY7 hepatitis a
28	107	100.0	56	12	091PB5 hepatitis a
29	107	100.0	56	12	091BQ0 hepatitis a
30	107	100.0	56	12	099HM1 hepatitis a
31	107	100.0	56	12	098VY0 hepatitis a
32	107	100.0	56	12	098VY8 hepatitis a
33	107	100.0	56	12	091PB1 hepatitis a
34	107	100.0	56	12	098VY6 hepatitis a
35	107	100.0	56	12	08JYK2 hepatitis a
36	107	100.0	56	12	098VY0 hepatitis a
37	107	100.0	56	12	08JYP7 hepatitis a
38	107	100.0	56	12	08JYP6 hepatitis a
39	107	100.0	56	12	091PB2 hepatitis a
40	107	100.0	56	12	091PB2 hepatitis a
41	107	100.0	56	12	091PB2 hepatitis a
42	107	100.0	56	12	098VY2 hepatitis a
43	107	100.0	56	12	098VY3 hepatitis a
44	107	100.0	56	12	067821 hepatitis a
45	107	100.0	56	12	091PB8 hepatitis a

## ALIGNMENTS

### RESULT 1

098VY2 ID 098VY2 PRELIMINARY; PRT; 53 AA.  
AC 098VY2;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE Polypeptide (fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID-12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Uruguay6;  
RA Costa-Matloff M., Ferre V., Monpocho S., Garcia L., Collina R.,  
RA Billand S., Vega I., Perez-Bercoff R., Cristina J.;  
RT "Genetic variability of Hepatitis A virus in South America reveals  
RT heterogeneity and co-circulation during epidemic outbreaks."  
RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.  
FT EMBL; AJ309232; CAC37078.1; -  
FT NON\_TER 1  
FT NON\_TER 53  
SQ SEQUENCE 53 AA; 6271 MW; 0846AF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;  
Best Local Similarity 100.0%; Pred. No. 4.2e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

098VY2 ID 098VY3 PRELIMINARY; PRT; 53 AA.  
AC 098VY3;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
DE Polypotein (Fragment).  
OC Hepatitis A virus.  
OS Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;  
OC Hepatovirus.  
NC NCB1\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Druil6;  
RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Collina R.,  
RT "Genetic variability of Hepatitis A virus in South America reveals  
heterogeneity and co-circulation during epidemic outbreaks.";  
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ309231; CAC37077.1; -.  
FT NON\_TER 1 53  
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;  
Query Match 100.0%; Score 107; DB 12; Length 53;  
Best Local Similarity 100.0%; Pred. No. 4.2e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKOR 20  
DB 29 SHIECRKPYKELRLEVGKOR 48  
|||||

RESULT 3  
Q98YV4 PRELIMINARY; PRT; 53 AA.  
ID Q98YV4  
AC Q98YV4;  
DT 01-JUN-2001 (TReMBLrel. 17, Created)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;  
OC Hepatovirus.  
NC NCB1\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Chile-J;  
RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Collina R.,  
RT "Genetic variability of Hepatitis A virus in South America reveals  
heterogeneity and co-circulation during epidemic outbreaks.";  
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ309230; CAC37076.1; -.  
FT NON\_TER 1 53  
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;  
Best Local Similarity 100.0%; Pred. No. 4.2e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKOR 20  
DB 29 SHIECRKPYKELRLEVGKOR 48  
|||||

RESULT 4  
Q98YV0 PRELIMINARY; PRT; 53 AA.  
ID Q98YV0  
AC Q98YV0;  
DT 01-JUN-2001 (TReMBLrel. 17, Created)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;  
OC Hepatovirus.  
NC NCB1\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Druil7;

OC Hepatovirus.  
OX NCB1\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Druinay4;  
RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Collina R.,  
RT "Genetic variability of Hepatitis A virus in South America reveals  
heterogeneity and co-circulation during epidemic outbreaks.";  
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ309234; CAC37080.1; -.  
FT NON\_TER 1 53  
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;  
Query Match 100.0%; Score 107; DB 12; Length 53;  
Best Local Similarity 100.0%; Pred. No. 4.2e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKOR 20  
DB 29 SHIECRKPYKELRLEVGKOR 48  
|||||

RESULT 5  
Q98YV5 PRELIMINARY; PRT; 53 AA.  
ID Q98YV5  
AC Q98YV5;  
DT 01-JUN-2001 (TReMBLrel. 17, Created)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;  
OC Hepatovirus.  
NC NCB1\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Druil3;  
RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Collina R.,  
RT "Genetic variability of hepatitis A virus in South America reveals  
heterogeneity and co-circulation during epidemic outbreaks.";  
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ309229; CAC37075.1; -.  
FT NON\_TER 1 53  
SQ SEQUENCE 53 AA; 6271 MW; 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;  
Best Local Similarity 100.0%; Pred. No. 4.2e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKOR 20  
DB 29 SHIECRKPYKELRLEVGKOR 48  
|||||

RESULT 6  
Q98YV6 PRELIMINARY; PRT; 53 AA.  
ID Q98YV6  
AC Q98YV6;  
DT 01-JUN-2001 (TReMBLrel. 17, Created)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;  
OC Hepatovirus.  
NC NCB1\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Druil7;

RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Colina R.,  
RA Billanduel S., Vega I., Perez-Bercoff R., Cristina J.;  
RT "Genetic variability of hepatitis A virus in South America reveals  
RT heterogeneity and co-circulation during epidemic outbreaks.";  
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ309227; CAC37074.1; -.  
FT NON\_TER 1  
FT NON\_TER 53  
SQ SEQUENCE 53 AA: 6271 MW: 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;  
Best Local Similarity 100.0%; Pred. No. 4.2e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELREVGKOR 20  
DB 29 SHIECRKPKYKELREVGKOR 48

## RESULT 7

O98VY1 PRELIMINARY; PRT; 53 AA.  
AC O98VY1:  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Chile16;  
RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Colina R.,  
RA Billanduel S., Vega I., Perez-Bercoff R., Cristina J.;  
RT "Genetic variability of Hepatitis A virus in South America reveals  
RT heterogeneity and co-circulation during epidemic outbreaks.";  
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ309233; CAC37079.1; -.  
FT NON\_TER 1  
FT NON\_TER 53  
SQ SEQUENCE 53 AA: 6271 MW: 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;  
Best Local Similarity 100.0%; Pred. No. 4.2e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SHIECRKPKYKELREVGKOR 20  
29 SHIECRKPKYKELREVGKOR 48

## RESULT 8

O98VY7 PRELIMINARY; PRT; 53 AA.  
AC O98VY7:  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Urui;  
RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Colina R.,  
RA Billanduel S., Vega I., Perez-Bercoff R., Cristina J.;  
RT "Genetic variability of hepatitis A virus in South America reveals  
RT heterogeneity and co-circulation during epidemic outbreaks.";  
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AJ309227; CAC37073.1; -.  
FT NON\_TER 1  
FT NON\_TER 53  
SQ SEQUENCE 53 AA: 6271 MW: 0846AEF4BC397432 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 53;  
Best Local Similarity 100.0%; Pred. No. 4.2e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELREVGKOR 20  
DB 29 SHIECRKPKYKELREVGKOR 48

## RESULT 9

O999T2 PRELIMINARY; PRT; 55 AA.  
AC O999T2:  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Urui3;  
RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Colina R.,  
RA Billanduel S., Vega I., Perez-Bercoff R., Cristina J.;  
RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and  
RT co-circulation during epidemic outbreaks.";  
RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ306386; CAC29235.1; -.  
DR InterPro: IPR000886; ER\_TARGET.  
DR PROSITE: PS00014; ER\_TARGET; 1.  
FT NON\_TER 1  
FT NON\_TER 55  
SQ SEQUENCE 55 AA: 6513 MW: 135CB05D46AEF4CB CRC64;

Query Match 100.0%; Score 107; DB 12; Length 55;  
Best Local Similarity 100.0%; Pred. No. 4.3e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELREVGKOR 20  
DB 28 SHIECRKPKYKELREVGKOR 47

## RESULT 10

O999T4 PRELIMINARY; PRT; 55 AA.  
AC O999T4:  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Chile-3;  
RA Costa-Mattoli M., Ferre V., Monpoeho S., Garcia L., Colina R.,  
RA Billanduel S., Vega I., Perez-Bercoff R., Cristina J.;  
RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and  
RT co-circulation during epidemic outbreaks.";  
RL Submitted (Feb-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ306384; CAC29233.1; -.  
DR InterPro: IPR000886; ER\_TARGET.  
DR PROSITE: PS00014; ER\_TARGET; 1.

FT NON\_TER 1 1  
FT SEQUENCE 55 55  
SQ SEQUENCE 55 AA; 6513 MW; 135CB05D46AEF4CB CRC64;

Query Match 100.0%; Score 107; DB 12; Length 55;  
Best Local Similarity 100.0%; Pred. No. 4.3e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELRLEVGKOR 20  
DB 28 SHIECRKPKYKELRLEVGKOR 47

## RESULT 11

0999T3 ID 0999T3 PRELIMINARY; PRT; 55 AA.  
AC 0999T3;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_Taxid=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Chile-9;  
RA Costa-Mattoli M., Ferre V., Monphoelo S., Garcia L., Collina R.,  
RT "Genetic Variability of Hepatitis C virus reveals heterogeneity and  
co-circulation during epidemic outbreaks."  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ306385; CAC29234.1; -  
DR InterPro; IPR000886; ER-target.  
DR PROSITE; PS00014; ER\_TARGET; 1.  
FT NON\_TER 1 1  
FT SEQUENCE 55 AA; 6513 MW; 135CB05D46AEF4CB CRC64;

Query Match 100.0%; Score 107; DB 12; Length 55;  
Best Local Similarity 100.0%; Pred. No. 4.3e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELRLEVGKOR 20  
DB 28 SHIECRKPKYKELRLEVGKOR 47

## RESULT 12

0999U8 ID 0999U8 PRELIMINARY; PRT; 55 AA.  
AC 0999U8;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_Taxid=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Arg-6;  
RA Billand S., Vega I., Perez-Bercoff R., Cristina J.;  
RT "Genetic Variability of Hepatitis A Virus in South America reveals  
heterogeneity and co-circulation during epidemic outbreaks."  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ306370; CAC29219.1; -  
DR InterPro; IPR000886; ER-target.  
DR PROSITE; PS00014; ER\_TARGET; 1.  
FT NON\_TER 1 1

FT NON\_TER 55 55  
FT SEQUENCE 55 AA; 6513 MW; 135CB05D46AEF4CB CRC64;

Query Match 100.0%; Score 107; DB 12; Length 55;  
Best Local Similarity 100.0%; Pred. No. 4.3e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELRLEVGKOR 20  
DB 28 SHIECRKPKYKELRLEVGKOR 47

## RESULT 13

Q8JYP4 ID Q8JYP4 PRELIMINARY; PRT; 56 AA.  
AC Q8JYP4;  
DT 01-OCT-2002 (TREMBlrel. 22, Created)  
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_Taxid=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Manuvelo;  
RA Theamboonlers A., Jantaradamee P., Poovorawan Y.;  
RT "Molecular characterization of Hepatitis A virus infection of an  
outbreak in the southern part of Thailand."  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF507065; AAM3425.1; -  
DR InterPro; IPR000886; ER-target.  
DR PROSITE; PS00014; ER\_TARGET; 1.  
FT NON\_TER 1 1  
FT SEQUENCE 56 AA; 6642 MW; 465CF51846AEF4BC CRC64;

Query Match 100.0%; Score 107; DB 12; Length 56;  
Best Local Similarity 100.0%; Pred. No. 4.4e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPKYKELRLEVGKOR 20  
DB 29 SHIECRKPKYKELRLEVGKOR 48

## RESULT 14

Q67822 ID Q67822 PRELIMINARY; PRT; 56 AA.  
AC Q67822;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_Taxid=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ISR-70;  
RA Karetnyi Y.V., Shulman L.M., Manor J., Lettner L., Shehab S.,  
RT "Genetic classification of hepatitis A virus strains isolated in  
Israel, based on their VP1/2A nucleotide sequence."  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z77247; CAB01040.1; -  
DR InterPro; IPR000886; ER-target.  
DR PROSITE; PS00014; ER\_TARGET; 1.  
FT NON\_TER 1 1  
FT CHAIN 1 >28 CAPSID PROTEIN VP1.  
FT CHAIN 29 >56 NONSTRUCTURAL PROTEIN NS2A.

FT NON\_TER 56 56  
 SO SEQUENCE 56 AA; 6614 MW; 8438C51846A6F4A4 CRC64;

Query Match 100.0%; Score 107; DB 12; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 4,4e-10;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLEVGR 20  
 |||||  
 Db 29 SHIECRKPYKELRLEVGR 48

## RESULT 15

039872 PRELIMINARY; PRT; 56 AA.

ID 039872  
 AC 039872;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)  
 DE Polypeptide (Fragment).

HE Hepatitis A virus.  
 Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 Hepatovirus.

OX NCBI\_TaxID=12092;

RN 11  
 RP SEQUENCE FROM N.A.

RC STRAIN=VDM;  
 RX MEDLINE=97247817; PubMed=9093940;

RA Taylor M.B.;  
 RT "Molecular epidemiology of South African strains of hepatitis A virus:  
 J. Med. Virol. 51:273-279(1997).

RL EMBL; U68697; AAB53593.1; -

DR InterPro; IPR000886; ER\_target.

DR PROSITE; PS00014; ER\_TARGET; 1.

FT NON\_TER 1 1

FT NON\_TER 56 56

SO SEQUENCE 56 AA; 6642 MW; 465CF51846A6F4BC CRC64;

Query Match 100.0%; Score 107; DB 12; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 4,4e-10;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLEVGR 20  
 |||||  
 Db 29 SHIECRKPYKELRLEVGR 48

Search completed: October 1, 2003, 10:02:34  
 Time : 35.1765 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 ; Search time 12.7059 Seconds  
(without alignments)  
66.600 Million cell updates/sec

Title: US-09-171-432a-39  
Perfect score: 107  
Sequence: 1 SHIECRKPKYKELRLEVGKOR 20

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 328717 seqs, 42310858 residues

al number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 10%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/Backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	1091	6	US-08-475-886-2 Patent No. 5516630-2
2	107	100.0	2227	3	US-08-475-886-2 Sequence 2, Appl
3	107	100.0	2227	3	US-08-475-886-4 Sequence 4, Appl
4	107	100.0	2227	3	US-08-475-886-6 Sequence 6, Appl
5	107	100.0	2227	3	US-08-397-233-2 Sequence 2, Appl
6	107	100.0	2227	3	US-08-397-233-4 Sequence 4, Appl
7	107	100.0	2227	3	US-09-171-387-2 Sequence 2, Appl
8	107	100.0	2227	4	US-09-653-499-2 Sequence 2, Appl
9	107	100.0	2227	4	US-09-653-499-4 Sequence 4, Appl
10	107	100.0	2227	4	US-09-653-499-6 Sequence 6, Appl
11	84.5	79.0	839	1	US-08-087-016-2 Sequence 2, Appl
12	43	40.2	1127	3	US-09-150-460B-11 Sequence 11, Appl
13	42.5	39.7	5405	3	US-08-718-388-9 Sequence 9, Appl
14	42	39.3	4444	4	US-09-252-991A-32415 Sequence 32415, A
15	42	39.3	472	4	US-09-328-352-6296 Sequence 6296, Ap
16	41	38.3	114	4	US-09-489-847-321 Sequence 321, App
17	41	38.3	391	5	PCT-US91-08177-3 Sequence 3, Appl
18	41	38.3	534	4	US-09-252-991A-32086 Sequence 32086, A
19	41	38.3	843	4	US-09-491-356C-20 Sequence 20, Appl
20	41	38.3	4544	1	US-08-469-486-52 Sequence 52, Appl
21	41	38.3	4544	2	US-08-469-486-52 Sequence 52, Appl
22	40	37.4	110	4	US-09-341-461-21 Sequence 21, Appl
23	40	37.4	414	4	US-09-252-991A-26659 Sequence 26659, A
24	39	36.4	24	1	US-08-406-347A-13 Sequence 13, Appl
25	39	36.4	154	4	US-09-198-452A-530 Sequence 530, App
26	39	36.4	238	4	US-09-257-179-80 Sequence 80, Appl
27	39	36.4	305	4	US-09-635-872A-3 Sequence 3, Appl

28	39	36.4	305	4	US-09-636-077A-3	Sequence 3, Appl
29	39	36.4	344	3	US-09-393-554-2	Sequence 2, Appl
30	39	36.4	495	4	US-09-252-991A-25802	Sequence 25802, A
31	39	36.4	520	6	5223391-5	Patent No. 5223391
32	39	36.4	892	4	US-09-585-858-16	Sequence 16, Appl
33	39	36.4	1124	4	US-09-252-991A-26810	Sequence 26810, A
34	39	36.4	1235	2	US-08-680-326-36	Sequence 36, Appl
35	38	35.5	31	1	US-08-015-770B-26	Sequence 26, Appl
36	38	35.5	31	1	US-08-488-252-15	Sequence 15, Appl
37	38	35.5	102	4	US-08-651-940-3	Sequence 3, Appl
38	38	35.5	102	4	US-09-295-029-3	Sequence 3, Appl
39	38	35.5	123	3	US-09-357-251-22	Sequence 22, Appl
40	38	35.5	133	4	US-09-252-991A-22809	Sequence 22809, A
41	38	35.5	155	2	US-08-651-940-4	Sequence 4, Appl
42	38	35.5	155	2	US-09-295-029-4	Sequence 4, Appl
43	38	35.5	256	2	US-08-719-758-2	Sequence 2, Appl
44	38	35.5	256	2	US-09-119-827-2	Sequence 2, Appl
45	38	35.5	341	4	US-09-198-452A-293	Sequence 293, App

## ALIGNMENTS

```
RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANIELLO, VINCENT R.;
; BAROUDY, BAHIGE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO: 2:
; LENGTH: 1091
5516630-2

Query Match      100.0%; Score 107; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 7.3e-10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SHIECRKPKYKELRLEVGKOR 20
Db      1029 SHIECRKPKYKELRLEVGKOR 1048

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262052
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
```

SEQ ID NO 2  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-08-475-886-2

Query Match 100.0%; Score 107; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.6e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKOR 20  
Db 792 SHIECRKPYKELRLEVGKOR 811

RESULT 3  
US-08-475-886-4  
Sequence 4, Application US/08475886A  
Patent No. 6113912  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 202642620S2  
CURRENT APPLICATION NUMBER: US/08/475,886A  
CURRENT FILING DATE: 1995-06-07  
EARLIER APPLICATION NUMBER: 07/947,338  
EARLIER FILING DATE: 1992-09-18  
EARLIER APPLICATION NUMBER: 08/397,232  
EARLIER FILING DATE: 1993-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
US-08-475-886-4

Query Match 100.0%; Score 107; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.6e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKOR 20  
Db 792 SHIECRKPYKELRLEVGKOR 811

RESULT 4  
US-08-475-886-6  
Sequence 6, Application US/08475886A  
Patent No. 6113912  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 202642620S2  
CURRENT APPLICATION NUMBER: US/08/475,886A  
CURRENT FILING DATE: 1995-06-07  
EARLIER APPLICATION NUMBER: 07/947,338  
EARLIER FILING DATE: 1992-09-18  
EARLIER APPLICATION NUMBER: 08/397,232  
EARLIER FILING DATE: 1993-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-08-475-886-6

Query Match 100.0%; Score 107; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.6e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKOR 20  
Db 792 SHIECRKPYKELRLEVGKOR 811

RESULT 5  
US-08-397-232-2  
Sequence 2, Application US/08397232A  
Patent No. 6180110  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 202642620S1  
CURRENT APPLICATION NUMBER: US/08/397,232A  
CURRENT FILING DATE: 1995-04-17  
EARLIER APPLICATION NUMBER: 07/947,338  
EARLIER FILING DATE: 1992-09-18  
EARLIER APPLICATION NUMBER: PCT/US93/08610  
EARLIER FILING DATE: 1993-09-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-08-397-232-2

Query Match 100.0%; Score 107; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.6e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLEVGKOR 20  
Db 792 SHIECRKPYKELRLEVGKOR 811

RESULT 6  
US-08-397-232-4  
Sequence 4, Application US/08397232A  
Patent No. 6180110  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 202642620S1  
CURRENT APPLICATION NUMBER: US/08/397,232A  
CURRENT FILING DATE: 1995-04-17  
EARLIER APPLICATION NUMBER: 07/947,338  
EARLIER FILING DATE: 1992-09-18  
EARLIER APPLICATION NUMBER: PCT/US93/08610  
EARLIER FILING DATE: 1993-09-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-08-397-232-4

Query Match 100.0%; Score 107; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.6e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLVGVKOR 20  
DB 792 SHIECRKPYKELRLVGVKOR 811

RESULT 7  
US-09-171-387-2  
Sequence 2, Application US/09171387  
Patent No. 6280734  
GENERAL INFORMATION:  
APPLICANT: RAYCHAUDHURI, GOPA;  
EMERSON, SUZANNE, U.;  
PURCELL, ROBERT, H.;  
TITLE OF INVENTION: SIMIAN-HUMAN HAV  
HAVING A CHIMERIC 2C PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MICROSOFT WORD 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/171,387  
FILING DATE: 24-Mar-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/06506  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US60/015,642  
FILING DATE: 19-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: William S. Feiler  
REGISTRATION NUMBER: 26,728  
REFERENCE/DOCKET NUMBER: 2026-42290S1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 2  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2227 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-171-387-2

Query Match 100.0%; Score 107; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.6e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLVGVKOR 20  
DB 792 SHIECRKPYKELRLVGVKOR 811  
RESULT 8  
US-09-653-499-2  
Sequence 2, Application US/09653499  
Patent No. 6423318  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
EMERSON, SUZANNE U  
PURCELL, ROBERT H  
D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262US2  
CURRENT APPLICATION NUMBER: US/09/653,499

CURRENT FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: 08/475,886  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-09-653-499-2

Query Match 100.0%; Score 107; DB 4; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.6e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLVGVKOR 20  
DB 792 SHIECRKPYKELRLVGVKOR 811

RESULT 9  
US-09-653-499-4  
Sequence 4, Application US/09653499  
Patent No. 6423318  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
EMERSON, SUZANNE U  
PURCELL, ROBERT H  
D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262US2  
CURRENT APPLICATION NUMBER: US/09/653,499  
CURRENT FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: 08/475,886  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
US-09-653-499-4

Query Match 100.0%; Score 107; DB 4; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.6e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SHIECRKPYKELRLVGVKOR 20  
DB 792 SHIECRKPYKELRLVGVKOR 811

RESULT 10  
US-09-653-499-6  
Sequence 6, Application US/09653499  
Patent No. 6423318  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
EMERSON, SUZANNE U  
PURCELL, ROBERT H  
D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262US2  
CURRENT APPLICATION NUMBER: US/09/653,499  
CURRENT FILING DATE: 2000-08-31

```
;; PRIOR APPLICATION NUMBER: 08/475,886
;; PRIOR FILING DATE: 1999-08-09
;; PRIOR APPLICATION NUMBER: 07/947,338
;; PRIOR FILING DATE: 1992-09-18
;; PRIOR APPLICATION NUMBER: 08/397,232
;; PRIOR FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO: 6
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match          100.0%; Score 107; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 SHIECRPKYKELRLEVGKOR 20
        |||||
        792 SHIECRPKYKELRLEVGKOR 811

RESULT 11
US-08-087-016-2
;; Sequence 2, Application US/08087016
;; Patent No. 5430135
;; GENERAL INFORMATION:
;; APPLICANT: NAINAN, OMANA V.
;; APPLICANT: MARGOLIS, HAROLD S.
;; APPLICANT: ROBERTSON, BETTY H.
;; APPLICANT: BRINTON, MARCO H.
;; APPLICANT: EBERT, JAMES W.
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
;; STREET: 1615 L Street N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: USA
;; ZIP: 20036
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentln Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/087,016
;; FILING DATE:
;; CLASSIFICATION: 424
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/678,828
;; FILING DATE: 03-APR-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: SCOTT, WATSON T.
;; REGISTRATION NUMBER: 26,581
;; REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 861-3000
;; TELEFAX: (202) 822-0944
;; TELEX: 6714627 CUSH
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 839 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-087-016-2

Query Match          79.0%; Score 84.5; DB 1; Length 839;
Best Local Similarity 90.0%; Pred. No. 3.8e-06;
Matches 18; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
```

```
QY      1 SHIECRPKYKELRLEVGKOR 20
        |||||
        792 SHIECRPKYKELRLEVGKOR 810

RESULT 12
US-09-150-460B-11
;; Sequence 11, Application US/09150460B
;; Patent No. 6190882
;; GENERAL INFORMATION:
;; APPLICANT: Lee, Cheng-Chi
;; APPLICANT: Albrecht, Urs
;; APPLICANT: Etchele, Gregor
;; APPLICANT: Sun, Zhong Sheng
;; TITLE OF INVENTION: Mammalian Circadian Rhythm-Like Gene
;; FILE REFERENCE: D6039
;; CURRENT APPLICATION NUMBER: US/09/150,460B
;; PRIOR FILING DATE: 1998-09-09
;; PRIOR APPLICATION NUMBER: US 60/058,256
;; PRIOR FILING DATE: 1997-09-09
;; NUMBER OF SEQ ID NOS: 21
;; SEQ ID NO: 11
;; LENGTH: 1127
;; TYPE: PRT
;; ORGANISM: D. melanogaster Period
;; FEATURE:
;; OTHER INFORMATION: Predicted protein sequence compared with RIGU1 4.7
US-09-150-460B-11

Query Match          40.2%; Score 43; DB 3; Length 1127;
Best Local Similarity 77.8%; Pred. No. 65;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 ECRPKYKEL 12
        |||||
        Db      392 ECRPKYKEL 400

RESULT 13
US-08-718-388-9
;; Sequence 9, Application US/08718388
;; Patent No. 6271362
;; GENERAL INFORMATION:
;; APPLICANT: MORIKAWA, MINORU
;; APPLICANT: HARADA, NAOKI
;; TITLE OF INVENTION: GENE ENCODING IgG Fc REGION-BINDING
;; TITLE OF INVENTION: PROTEIN
;; NUMBER OF SEQUENCES: 29
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH
;; STREET: PO BOX 747
;; CITY: FALLS CHURCH
;; STATE: VA
;; COUNTRY: USA
;; ZIP: 22040-0747
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentln Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/718,388
;; FILING DATE:
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: MURPHY JR, GERALD M.
;; REGISTRATION NUMBER: 28,977
;; REFERENCE/DOCKET NUMBER: 0230-111
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 205-8000
;; TELEFAX: (703) 205-8050
;; INFORMATION FOR SEQ ID NO: 9:
```

SEQUENCE CHARACTERISTICS:  
 LENGTH: 5405 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-718-388-9

Query Match 39.7%; Score 42.5; DB 3; Length 5405;  
 Best Local Similarity 39.1%; Pred. No. 4.4e+02;  
 Matches 9; Conservative 5; Mismatches 4; Indels 5; Gaps 1;

OY 2 HIECRK----PYKELREVGKQ 19  
 DB 4824 HVTCEGAGACGPHCEKLEEDGVO 4846

RESULT 14  
 US-09-252-991A-32415  
 Sequence 32415, Application US/09252991A  
 Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

PRIOR FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 32415

LENGTH: 444

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-32415

Query Match 39.3%; Score 42; DB 4; Length 444;  
 Best Local Similarity 50.0%; Pred. No. 34;  
 Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 3 IECKRPYKELLEV 16  
 DB 135 VOCRLPHRALRLGV 148

RESULT 15

US-09-328-352-6296

Sequence 6296, Application US/09328352

Patent No. 6562958

GENERAL INFORMATION:

APPLICANT: Gary L. Breton et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

FILE REFERENCE: GTC99-03PA

CURRENT APPLICATION NUMBER: US/09/328,352

CURRENT FILING DATE: 1999-06-04

NUMBER OF SEQ ID NOS: 8252

SEQ ID NO 6296

LENGTH: 472

TYPE: PRT

ORGANISM: Acinetobacter baumannii

US-09-328-352-6296

Query Match 39.3%; Score 42; DB 4; Length 472;  
 Best Local Similarity 40.9%; Pred. No. 37;  
 Matches 9; Conservative 8; Mismatches 3; Indels 2; Gaps 1;

OY 1 SHIECRKPYK--ELREVGKOR 20  
 DB 211 SHLENQRLKQOVDLKLQYAKQO 232

Search completed: October 1, 2003, 10:06:28  
 Job time : 14.7059 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993-2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 ; Search time: 24.1176 seconds  
(without alignments)  
131.201 Million cell updates/sec

Title: US-09-171-432a-39  
Perfect score: 107  
Sequence: 1 SHIECRKPKYKELRLEVGKOR 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	352	15 US-10-272-459-45	Sequence 45, Appl
2	107	100.0	836	15 US-10-272-459-40	Sequence 41, Appl
3	107	100.0	980	15 US-10-272-459-41	Sequence 40, Appl
4	107	100.0	2227	14 US-09-929-955-12	Sequence 12, Appl
5	107	100.0	2227	14 US-10-104-966-12	Sequence 12, Appl
6	107	100.0	2227	14 US-10-135-988-2	Sequence 2, Appl
7	107	100.0	2227	14 US-10-135-988-4	Sequence 4, Appl
8	107	100.0	2227	14 US-10-135-988-6	Sequence 6, Appl
9	45.5	42.5	178	12 US-10-091-007-124	Sequence 124, App
10	43	40.2	45	9 US-09-728-912-13	Sequence 13, Appl
11	43	40.2	52	9 US-09-728-912-12	Sequence 12, Appl
12	43	40.2	413	9 US-09-728-912-2	Sequence 2, Appl
13	42.5	39.7	5405	14 US-09-922-217-1116	Sequence 1116, Ap
14	42.5	39.7	5405	14 US-10-025-380-1116	Sequence 1116, Ap
15	42	39.3	1289	10 US-09-933-145-11	Sequence 11, Appl

## ALIGNMENTS

16	41	38.3	22	12	US-10-173-461-21	Sequence 21, Appl
17	41	38.3	39	12	US-10-173-461-65	Sequence 65, Appl
18	41	38.3	71	9	US-09-864-761-43211	Sequence 43211, A
19	41	38.3	80	9	US-09-804-156-38	Sequence 38, Appl
20	41	38.3	80	12	US-10-319-519-38	Sequence 38, Appl
21	41	38.3	80	14	US-10-067-761-38	Sequence 38, Appl
22	41	38.3	153	11	US-09-533-029-96	Sequence 96, Appl
23	41	38.3	153	11	US-09-934-445-40	Sequence 40, Appl
24	41	38.3	153	15	US-10-286-264-108	Sequence 108, App
25	41	38.3	153	16	US-10-278-536-126	Sequence 126, App
26	41	38.3	254	9	US-09-820-893-94	Sequence 94, Appl
27	41	38.3	264	12	US-10-173-461-2	Sequence 2, Appl
28	41	38.3	315	9	US-09-820-893-59	Sequence 59, Appl
29	41	38.3	324	9	US-09-820-893-96	Sequence 96, Appl
30	41	38.3	492	10	US-09-978-295A-7	Sequence 7, Appl
31	41	38.3	492	10	US-09-978-697-7	Sequence 7, Appl
32	41	38.3	492	10	US-09-978-832A-7	Sequence 7, Appl
33	41	38.3	492	10	US-09-999-832A-7	Sequence 7, Appl
34	41	38.3	492	11	US-09-978-189-7	Sequence 7, Appl
35	41	38.3	492	11	US-09-978-608A-7	Sequence 7, Appl
36	41	38.3	492	11	US-09-978-585A-7	Sequence 7, Appl
37	41	38.3	492	11	US-09-978-191A-7	Sequence 7, Appl
38	41	38.3	492	11	US-09-978-403A-7	Sequence 7, Appl
39	41	38.3	492	11	US-09-978-564A-7	Sequence 7, Appl
40	41	38.3	492	11	US-09-999-833A-7	Sequence 7, Appl
41	41	38.3	492	11	US-09-981-915A-7	Sequence 7, Appl
42	41	38.3	492	11	US-09-978-824-7	Sequence 7, Appl
43	41	38.3	492	11	US-09-918-585A-7	Sequence 7, Appl
44	41	38.3	492	11	US-09-978-433A-7	Sequence 7, Appl
45	41	38.3	492	11	US-09-978-193A-7	Sequence 7, Appl

RESULT 1  
US-10-272-459-45  
Sequence 45, Application US/10272459  
Publication No. US20030124517A1  
GENERAL INFORMATION:  
APPLICANT: PICHANTRES, Sergio  
TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT  
FILE REFERENCE: PP17955.002 / 2301-17955  
CURRENT APPLICATION NUMBER: US/10/272,459  
CURRENT FILING DATE: 2002-10-15  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 45  
LENGTH: 352  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: recombinant protein  
US-10-272-459-45

Query Match 100.0% Score 107; DB 15; Length 352;  
Best Local Similarity 100.0% Pred. No. 3.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 SHIECRKPKYKELRLEVGKOR 20  
Db 164 SHIECRKPKYKELRLEVGKOR 183

RESULT 2  
US-10-272-459-40  
Sequence 40, Application US/10272459  
Publication No. US20030124517A1  
GENERAL INFORMATION:  
APPLICANT: PICHANTRES, Sergio  
TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

;; TITLE OF INVENTION: PROTEINS AND USES THEREOF  
;; FILE REFERENCE: P17955.002 / 2301-17955  
;; CURRENT APPLICATION NUMBER: US/10/272.459  
;; CURRENT FILING DATE: 2002-10-15  
;; NUMBER OF SEQ ID NOS: 50  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 40  
;; LENGTH: 836  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein  
US-10-272-459-40

Query Match 100.0%; Score 107; DB 15; Length 836;  
Best Local Similarity 100.0%; Pred. No. 8.7e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRPKYKELRLEVGKOR 20  
|||||  
792 SHIECRPKYKELRLEVGKOR 811

RESULT 3  
US-10-272-459-41  
;; Sequence 41, Application US/10272459  
;; Publication No. US20030124517A1  
;; GENERAL INFORMATION:  
;; APPLICANT: PICHAUNTES, Sergio  
;; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT  
;; TITLE OF INVENTION: PROTEINS AND USES THEREOF  
;; FILE REFERENCE: P17955.002 / 2301-17955  
;; CURRENT APPLICATION NUMBER: US/10/272.459  
;; CURRENT FILING DATE: 2002-10-15  
;; NUMBER OF SEQ ID NOS: 50  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 41  
;; LENGTH: 980  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein  
US-10-272-459-41

Query Match 100.0%; Score 107; DB 15; Length 980;  
Best Local Similarity 100.0%; Pred. No. 1e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRPKYKELRLEVGKOR 20  
|||||  
792 SHIECRPKYKELRLEVGKOR 811

RESULT 4  
US-09-929-955-12  
;; Sequence 12, Application US/09929955  
;; Patent No. US20020136740A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Matti Salberg  
;; APPLICANT: Catharina Hultgren  
;; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND  
;; TITLE OF INVENTION: METHODS OF USE THEREOF  
;; FILE REFERENCE: TRIPEP.23AUS2  
;; CURRENT APPLICATION NUMBER: US/09/929.955  
;; CURRENT FILING DATE: 2001-08-15  
;; PRIOR APPLICATION NUMBER: 09/705,547  
;; PRIOR FILING DATE: 2000-11-03  
;; PRIOR APPLICATION NUMBER: 60/229,175  
;; PRIOR FILING DATE: 2000-08-29  
;; PRIOR APPLICATION NUMBER: 60/225,767  
;; PRIOR FILING DATE: 2000-08-17

;; NUMBER OF SEQ ID NOS: 49  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 12  
;; LENGTH: 2227  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Hepatitis A virus sequence  
US-09-929-955-12

Query Match 100.0%; Score 107; DB 10; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.5e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRPKYKELRLEVGKOR 20  
|||||  
Db 792 SHIECRPKYKELRLEVGKOR 811

RESULT 5  
US-10-104-966-12  
;; Sequence 12, Application US/10104966  
;; Publication No. US20020155124A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Matti Salberg  
;; APPLICANT: Catharina Hultgren  
;; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND  
;; TITLE OF INVENTION: METHODS OF USE THEREOF  
;; FILE REFERENCE: TRIPEP.23AUSC1  
;; CURRENT APPLICATION NUMBER: US/10/104.966  
;; CURRENT FILING DATE: 2002-03-22  
;; PRIOR APPLICATION NUMBER: 09/705,547  
;; PRIOR FILING DATE: 2000-11-03  
;; PRIOR APPLICATION NUMBER: 60/229,175  
;; PRIOR FILING DATE: 2000-08-29  
;; NUMBER OF SEQ ID NOS: 15  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 12  
;; LENGTH: 2227  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Hepatitis A virus sequence  
US-10-104-966-12

Query Match 100.0%; Score 107; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.5e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRPKYKELRLEVGKOR 20  
|||||  
Db 792 SHIECRPKYKELRLEVGKOR 811

RESULT 6  
US-10-135-988-2  
;; Sequence 2, Application US/10135988  
;; Publication No. US20020176869A1  
;; GENERAL INFORMATION:  
;; APPLICANT: FUNKHOUSER, ANN W  
;; APPLICANT: EMERSON, SUZANNE U  
;; APPLICANT: PORCELL, ROBERT H  
;; APPLICANT: D'HONDT, ERIC  
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
;; FILE REFERENCE: 20264262US3  
;; CURRENT APPLICATION NUMBER: US/10/135.988  
;; CURRENT FILING DATE: 2002-04-29  
;; PRIOR APPLICATION NUMBER: 07/947,338  
;; PRIOR FILING DATE: 1992-09-18  
;; PRIOR APPLICATION NUMBER: 08/397,232  
;; PRIOR FILING DATE: 1995-03-10  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: PatentIn Ver. 2.1



SEQ ID NO 2  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-10-135-988-2

Query Match 100.0%; Score 107; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.5e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLEVGRKOR 20  
DB 792 SHIECRKPYKELRLEVGRKOR 811

RESULT 7  
US-10-135-988-4

Sequence 4, Application US/10135988  
Publication No. US20020176869A1

GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W

APPLICANT: PURCELL, ROBERT H

TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

FILE REFERENCE: 20264262US3

CURRENT FILING DATE: 2002-04-29

PRIOR FILING DATE: 1992-09-18

PRIOR APPLICATION NUMBER: 08/397,232

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 4

LENGTH: 2227

TYPE: PRT

ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
US-10-135-988-4

Query Match 100.0%; Score 107; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.5e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLEVGRKOR 20  
DB 792 SHIECRKPYKELRLEVGRKOR 811

RESULT 8

US-10-135-988-6

Sequence 6, Application US/10135988  
Publication No. US20020176869A1

GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W

APPLICANT: EMERSON, SUZANNE U

APPLICANT: PURCELL, ROBERT H

TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

FILE REFERENCE: 20264262US3

CURRENT FILING DATE: 2002-04-29

PRIOR FILING DATE: 1992-09-18

PRIOR APPLICATION NUMBER: 08/397,232

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 6

LENGTH: 2227

TYPE: PRT

ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-10-135-988-6

Query Match 100.0%; Score 107; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.5e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SHIECRKPYKELRLEVGRKOR 20  
DB 792 SHIECRKPYKELRLEVGRKOR 811

RESULT 9  
US-10-091-007-124

Sequence 124, Application US/10091007

Publication No. US20030170782A1

GENERAL INFORMATION:

APPLICANT: Microbial Technics Limited

APPLICANT: Le Page, Richard W F Wells, Jeremy M

TITLE OF INVENTION: Hamifly, Sean B

FILE REFERENCE: PMC/P21978MO

CURRENT FILING DATE: 2002-03-06

PRIOR FILING DATE: 1999-09-07

NUMBER OF SEQ ID NOS: 276

SOFTWARE: PatentIn version 3.0

SEQ ID NO 124

LENGTH: 178

TYPE: PRT

ORGANISM: Streptococcus agalactiae  
US-10-091-007-124

Query Match 42.5%; Score 45.5; DB 12; Length 178;  
Best Local Similarity 38.5%; Pred. No. 14;  
Matches 10; Conservative 4; Mismatches 3; Indels 9; Gaps 1;

OY 3 IECKR-----PYKELRLEVGRKQ 19  
DB 103 IKCKIKMAMLRVFPFKELRLSAGRE 128

RESULT 10  
US-09-728-912-13

Sequence 13, Application US/09728912

Patent No. US2001003643A1

GENERAL INFORMATION:

APPLICANT: Holloway, James L.

TITLE OF INVENTION: Tumor Necrosis Factor-Stimulated Gene

FILE REFERENCE: 99-94US

CURRENT FILING DATE: 2000-12-01

PRIOR FILING DATE: 1999-12-06

NUMBER OF SEQ ID NOS: 14

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 13

LENGTH: 45

TYPE: PRT

ORGANISM: Homo sapiens  
US-09-728-912-13

Query Match 40.2%; Score 43; DB 9; Length 45;  
Best Local Similarity 53.3%; Pred. No. 7.9;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 5 CRKPYKELRLEVGRKQ 19  
DB 9 CRNPTRRLRVREGVQ 23

RESULT 11  
US-09-728-912-12

```
; Sequence 12, Application US/09728912
; Patent No. US20010036643A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Tumor Necrosis Factor-Stimulated Gene
; FILE REFERENCE: 99-94US
; CURRENT APPLICATION NUMBER: US/09/728,912
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/169,252
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 52
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-728-912-12
```

```
Query Match          40.2%; Score 43; DB 9; Length 52;
Best Local Similarity 53.3%; Pred. No. 9.2;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 5 CRPKYKELLEVQKQ 19
    |||:||||:|
Db 28 CRNPTRELRYREGVQ 42
```

```
RESULT 12
US-09-728-912-2
; Sequence 2, Application US/09728912
; Patent No. US20010036643A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: Tumor Necrosis Factor-Stimulated Gene
; FILE REFERENCE: 99-94US
; CURRENT APPLICATION NUMBER: US/09/728,912
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/169,252
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 413
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-728-912-2
```

```
Query Match          40.2%; Score 43; DB 9; Length 413;
Best Local Similarity 53.3%; Pred. No. 85;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 5 CRPKYKELLEVQKQ 19
    |||:||||:|
Db 376 CRNPTRELRYREGVQ 390
```

```
RESULT 13
US-09-922-217-1116
; Sequence 1116, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiaqichun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Mesgher, Madeleine Joy
; APPLICANT: Stoik, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
```

```
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1116
; LENGTH: 5405
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-922-217-1116
```

```
Query Match          39.7%; Score 42.5; DB 9; Length 5405;
Best Local Similarity 39.1%; Pred. No. 1.6e+03;
Matches 9; Conservative 5; Mismatches 4; Indels 5; Gaps 1;
```

```
QY 2 HIECRK----PYKELLEVQKQ 19
    |||:||||:|
Db 4824 HVTQCGACGACGPHECRLEDGVQ 4846
```

```
RESULT 14
US-10-025-380-1116
; Sequence 1116, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiaqichun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Mesgher, Madeleine Joy
; APPLICANT: Stoik, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darick
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025,380
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1116
; LENGTH: 5405
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-025-380-1116
```

```
Query Match          39.7%; Score 42.5; DB 14; Length 5405;
Best Local Similarity 39.1%; Pred. No. 1.6e+03;
Matches 9; Conservative 5; Mismatches 4; Indels 5; Gaps 1;
```

```
QY 2 HIECRK----PYKELLEVQKQ 19
    |||:||||:|
Db 4824 HVTQCGACGACGPHECRLEDGVQ 4846
```

```
RESULT 15
US-09-932-145-11
; Sequence 11, Application US/09932145
; Patent No. US20020161191A1
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
```

APPLICANT: Mintier, Gabe  
APPLICANT: Kinney, Gene G  
APPLICANT: Ramnathan, Chandra S  
TITLE OF INVENTION: NOVEL IMIDAZOLINE RECEPTOR HOMOLOGS  
FILE REFERENCE: D0020 NP  
CURRENT APPLICATION NUMBER: US/09/932,145  
CURRENT FILING DATE: 2001-08-17  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 11  
LENGTH: 1289  
TYPE: prt  
ORGANISM: Drosophila melanogaster  
FEATURE:  
OTHER INFORMATION: DATABASE ACCESSION NUMBER: Genbank,  
OTHER INFORMATION: Accession: AAF52305  
US-09-932-145-11

Query Match 39.3%; Score 42; DB 10; Length 1289;  
Best Local Similarity 42.1%; Pred. No. 4.2e+02;  
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
1 SHIECKRPYKELRLVSKQ 19  
||:|:|:|:|:|:|:|:|:|:  
DB 442 SHLETKKIEPLRLTYGNE 460

Search completed: October 1, 2003, 10:37:49  
Job time : 24.1176 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

..OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 12.8235 Seconds  
(without alignments)  
149.988 Million cell updates/sec

Title: US-09-171-432a-42

Perfect score: 101

Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_76:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	852	1 GNNYHA	genome polyprotein
2	101	100.0	1358	2 A03905	genome polyprotein
3	101	100.0	2227	1 GNNYHM	genome polyprotein
4	101	100.0	2227	1 GNNYHR	genome polyprotein
5	101	100.0	2227	1 GNNYMR	genome polyprotein
6	96	95.0	2230	1 GNNYSA	genome polyprotein
7	95	94.1	2227	1 GNNYHB	genome polyprotein
8	84	83.2	839	1 GNNYS2	genome polyprotein
9	75	74.3	341	2 S04137	genome polyprotein
10	52.5	52.0	1025	2 A54718	genome polyprotein
11	47	46.5	414	2 D96838	genome polyprotein
12	45	44.6	79	2 AG1062	genome polyprotein
13	45	44.6	321	2 E86423	genome polyprotein
14	44	43.6	340	1 MMES7	genome polyprotein
15	44	43.6	343	2 S74937	genome polyprotein
16	44	43.6	363	2 S77299	genome polyprotein
17	44	43.6	775	2 T48957	genome polyprotein
18	43	42.6	189	2 A12534	genome polyprotein
19	43	42.6	309	2 D75008	genome polyprotein
20	43	42.6	379	2 T49919	genome polyprotein
21	43	42.6	389	2 S41748	genome polyprotein
22	42.5	42.1	105	2 B12398	genome polyprotein
23	42	41.6	109	2 S69307	genome polyprotein
24	42	41.6	165	2 S02053	genome polyprotein
25	42	41.6	178	2 T02335	genome polyprotein
26	42	41.6	834	2 F82673	genome polyprotein
27	41.5	41.1	304	2 H82984	genome polyprotein
28	41	40.6	83	2 T17809	genome polyprotein
29	41	40.6	159	2 C72210	genome polyprotein

30	41	40.6	213	2 H86638	hypothetical prote
31	41	40.6	229	2 T19153	hypothetical prote
32	41	40.6	237	2 C84065	transcription regu
33	41	40.6	370	2 G97001	endoglucanase faml
34	41	40.6	561	2 G36752	unknown protein f2
35	41	40.6	622	2 T37257	hypothetical prote
36	41	40.6	897	2 T06540	polyribonucleotide
37	41	40.6	1040	2 A34695	axonal glycoprotein
38	41	40.6	1234	2 T30254	junction protein -
39	41	40.1	932	2 T25007	hypothetical prote
40	40.5	40.1	932	2 T25008	hypothetical prote
41	40	39.6	177	2 C64133	mercuric resistanc
42	40	39.6	238	2 C83240	probable transcrip
43	40	39.6	259	2 F64532	conserved hypothet
44	40	39.6	272	2 T49070	probable heat shoc
45	40	39.6	292	2 S23239	hypothetical prote

#### ALIGNMENTS

RESULT 1

GNNYHA

genome polyprotein - human hepatitis A virus (strain C8326) (fragment)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C>Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Jul-1999

C:Accession: A03904

R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, J. Virol. 54, 247-255, 1985

A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.

A:Reference number: A03904; MUID:85185648; PMID:2985793

A:Accession: A03904

A:Molecule type: genomic RNA

A:Residues: 1-852 <LIN>

A:Cross-references: EMBL:M10033; NID:q329592; PIDN:AAA5470.1; PID:q329593

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; polyprotein

F:1-245/Product: coat protein 1A #status predicted <C1A>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 100.0% Score 101; DB 1; Length 852;

Best Local Similarity 100.0%; Pred. NO. 1.2e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20

DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 2

A03905

genome polyprotein (version 2) - human hepatitis A virus (fragments)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; prote

C:Species: human hepatitis A virus

C>Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 15-Nov-1996

C:Accession: A03905

R:Baroudy, B.M.; Ticehurst, J.R.; Mele, T.A.; Malzel Jr., J.V.; Purcell, R.H.; Feins

A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and R

A:Reference number: A03905; MUID:85166289; PMID:2984684

A:Accession: A03905

A:Molecule type: genomic RNA

A:Residues: 1-1358 <BAR>

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; polyprotein

F:1-245/Product: coat protein 1A #status predicted <C1A>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-854/Product: core protein 2A (fragment) #status predicted <C2A>



1 Gen. Virol. 72, 1677-1683, 1991  
 A>Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and  
 A:Reference number: J01080; MUID:91311420; PMID:1649901  
 A:Contents: annotation  
 A>Note: neither amino acid nor nucleotide sequence is given  
 R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhnik  
 submitted to the EMBL Data Library, May 1989  
 A:Reference number: S04885  
 A:Accession: S04885  
 A:Molecule type: genomic RNA  
 A:Residues: 1750-2164 <BAL1>  
 A:Cross-references: EMBL:X15461; NID:961971; PIDN:CAA33490.1; PID:g930268  
 R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhnik  
 FBS Lett. 247, 425-428, 1989  
 A>Title: Variations in genome fragments coding for RNA polymerase in human and simian he  
 A:Reference number: S03965; MUID:89232168; PMID:2541023  
 A:Accession: S03965  
 A:Molecule type: genomic RNA  
 A:Residues: 1960-2164 <BAL2>  
 A:Cross-references: EMBL:X15461  
 C:Superfamily: hepatitis A virus genome polyprotein  
 C:Keywords: coat protein; core protein; polyprotein  
 F:1-27/Product: coat protein 1A #status predicted <C1A>  
 F:28-249/Product: coat protein 1B #status predicted <C1B>  
 F:250-495/Product: coat protein 1C #status predicted <C1C>  
 F:496-795/Product: coat protein 1D #status predicted <C1D>  
 F:796-984/Product: coat protein 2A #status predicted <C2A>  
 F:985-1091/Product: core protein 2B #status predicted <C2B>  
 F:1092-1426/Product: core protein 2C #status predicted <C2C>  
 F:1427-1498/Product: protein 3A #status predicted <P3A>  
 F:1499-1521/Product: protein 3B #status predicted <P3B>  
 F:1522-1741/Product: protein 3C #status predicted <P3C>  
 F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 95.0%; Score 96; DB 1; Length 2230;  
 Best Local Similarity 90.0%; Pred. No. 2.5e-07;  
 Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPRRKMKGLFSQAKISLF 20  
 DB 827 ILPPRRKMKGLFSQAKISLF 846

RESULT 7  
 GNNYHB  
 genome polyprotein - human hepatitis A virus (strain MBB)  
 N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro  
 Vpg: protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48); protein 3D  
 Species: human hepatitis A virus  
 Note: host Homo sapiens (man)  
 C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 10-May-1996  
 C:Accession: J50303  
 R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Klehn, R.; Wimmer, E.; Deinhardt,  
 Virus Res. 8, 153-171, 1987  
 A>Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat  
 A:Reference number: J50303; MUID:88045071; PMID:2823500  
 A:Accession: J50303  
 A:Molecule type: genomic RNA  
 A:Residues: 1-2227 <PAU>  
 A:Cross-references: EMBL:M20273  
 C:Superfamily: hepatitis A virus genome polyprotein  
 C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydr  
 F:1-23/Product: coat protein 1A #status predicted <VP4>  
 F:24-246/Product: coat protein 1B #status predicted <VP3>  
 F:247-491/Product: coat protein 1C #status predicted <VP2>  
 F:492-836/Product: coat protein 1D #status predicted <VP1>  
 F:837-980/Product: core protein 2A #status predicted <P2A>  
 F:981-1108/Product: core protein 2B #status predicted <P2B>  
 F:1109-1438/Product: core protein 2C #status predicted <P2C>  
 F:1439-1496/Product: protein 3A #status predicted <P3A>  
 F:1497-1519/Product: genome-linked protein Vpg #status predicted <VP6>  
 F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>  
 F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 94.1%; Score 95; DB 1; Length 2227;  
 Best Local Similarity 95.0%; Pred. No. 3.6e-07;  
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLPPRRKMKGLFSQAKISLF 20  
 DB 823 VLPPRRKMKGLFSQAKISLF 842

RESULT 8  
 GNNYS2  
 genome polyprotein - simian hepatitis A virus (strain CY-145) (fragment)  
 N:Contains: amino end of core protein 2A; coat protein 1A; coat protein 1B; coat prot  
 C:Species: simian hepatitis A virus  
 A>Note: host Macaca fascicularis (cynomolgus macaque)  
 C:Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 16-Jul-1999  
 C:Accession: J01180  
 R:Naiman, O.V.; Margolis, H.S.; Robertson, B.H.; Balayan, M.; Brinton, M.A.  
 J. Gen. Virol. 72, 1685-1689, 1991  
 A>Title: Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus  
 A:Reference number: J01180; MUID:91311421; PMID:1649902  
 A:Accession: J01180  
 A:Molecule type: genomic RNA  
 A:Residues: 1-839 <NA1>  
 A:Cross-references: GB:M59286; NID:9329599; PIDN:AAA5473.1; PID:g555083  
 C:Superfamily: hepatitis A virus genome polyprotein  
 C:Keywords: coat protein; core protein; glycoprotein; polyprotein  
 F:1-23/Product: coat protein 1A #status predicted <VP0>  
 F:24-245/Product: coat protein 1B #status predicted <VP3>  
 F:246-491/Product: coat protein 1C #status predicted <VP1>  
 F:492-839/Product: core protein 2A (fragment) #status predicted <P2>  
 F:261,312,728,756/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.2%; Score 84; DB 1; Length 839;  
 Best Local Similarity 83.3%; Pred. No. 8.6e-06;  
 Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPRRKMKGLFSQAKIS 18  
 DB 822 ILPPRRKMKGLFSQAKIS 839

RESULT 9  
 S04137  
 genome polyprotein - human hepatitis A virus (strain LCDC-1) (fragment)  
 C:Species: human hepatitis A virus  
 C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 21-Jul-2000  
 C:Accession: S04137  
 R:Andonov, A.P.; Lau, P.C.K.; Chaudhary, R.  
 Nucleic Acids Res. 17, 3594, 1989  
 A>Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A vir  
 A:Reference number: S04137; MUID:89263805; PMID:2542903  
 A:Accession: S04137  
 A:Molecule type: mRNA  
 A:Residues: 1-341 <AND>  
 A:Cross-references: EMBL:X14666; NID:962301; PIDN:CAA32794.1; PID:g4377576  
 C:Genetics:  
 A:Gene: VP1  
 C:Superfamily: hepatitis A virus genome polyprotein  
 C:Keywords: coat protein; polyprotein  
 F:2-340/Product: coat protein 1D (VP1) #status predicted <MAT>  
 Query Match 74.3%; Score 75; DB 2; Length 341;  
 Best Local Similarity 93.3%; Pred. No. 0.0001;  
 Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPRRKMKGLFSQA 15  
 DB 327 VLPPRRKMKGLFSQS 341

RESULT 10

A54718  
 dihydropyrimidine dehydrogenase (NADP) (EC 1.3.1.2) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 28-Apr-1995 #sequence\_revision 28-Apr-1995 #text\_change 03-Jun-2002  
 C:Accession: A54718  
 R:Yokota, H.; Fernandez-Salguero, P.; Furuya, H.; Lin, K.; McBride, O.W.; Podschun, B.;  
 J. Biol. Chem. 269, 23192-23196, 1994  
 A:Title: CDNA cloning and chromosome mapping of human dihydropyrimidine dehydrogenase,  
 A:Reference number: A54718; MUID:94365020; PMID:8083224  
 A:Accession: A54718  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 11025 <OK>  
 A:Cross-references: GB:009178; NID:9558304; PID:9558305  
 A:Note: authors translated the codon GGC for residue 748 as Asp, and TTA for residue 101  
 C:Genetics:  
 A:Gene: GDB:DPYD  
 A:Cross-references: GDB:364102; OMIM:274270  
 A:Map position: 1p22-1p22  
 C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] h  
 C:Keywords: 4Fe-4S; flavoprotein; homodimer; iron-sulfur protein; NADP; oxidoreductase  
 46-1005/Domain: ferredoxin 2[4Fe-4S] homology <FER>

Query Match 52.0%; Score 52.5; DB 2; Length 1025;  
 Best Local Similarity 63.2%; Pred. No. 2;  
 Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

OY 2 LPPPRKMKGLFSQAKISLF 20  
 ||||| : |||||  
 Db 176 LPPPRKMKSEAVS-AKIALF 193

RESULT 11  
 D96838  
 unknown protein T21F11.5 [Imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
 C:Accession: D96838  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
 ansen, N.F.; Hughes, B.; Hulzar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzalla,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: D96838  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-414 <STO>  
 A:Cross-references: GB:AE005172; NID:96730725; PIDN:AAF27115.1; GSPDB:GN00141  
 C:Genetics:  
 A:Gene: T21F11.5  
 A:Map position: 1

Query Match 46.5%; Score 47; DB 2; Length 414;  
 Best Local Similarity 56.2%; Pred. No. 6.2;  
 Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 3 PPRKMKGLFSQAKIS 18  
 ||||| : |||||  
 Db 106 PPRMDQLFRQNVLS 121

RESULT 12  
 AG1062  
 bacteriophage gene regulatory protein STY4826 [Imported] - Salmonella enterica subsp. en  
 C:Species: Salmonella enterica subsp. enterica serovar Typh  
 A:Note: this species has also been called Salmonella typhi  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C:Accession: AG1062  
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church  
 th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
 S.; Moulé, S.; O'Garra, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se  
 A:Reference number: AB0502; MUID:21534947; PMID:11677608  
 A:Accession: AG1062  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 179 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD06948.1; PID:916505595; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: STY4826

Query Match 44.6%; Score 45; DB 2; Length 79;  
 Best Local Similarity 63.6%; Pred. No. 2.1;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 3 PPRKMKGLFS 13  
 ||||| : |||||  
 Db 64 PPRKMGCCYS 74

RESULT 13  
 E86423  
 probable 60S ribosomal protein L18A - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 27-Nov-2001  
 C:Accession: E86423  
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
 ansen, N.F.; Hughes, B.; Hulzar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marzalla,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shin, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID:21016719; PMID:11130712  
 A:Accession: E86423  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-321 <STO>  
 A:Cross-references: GB:AE005172; NID:910092460; PIDN:AA12862.1; GSPDB:GN00141  
 C:Genetics:  
 A:Map position: 1

Query Match 44.6%; Score 45; DB 2; Length 321;  
 Best Local Similarity 50.0%; Pred. No. 10;  
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 1 VLPFRKMKGLFSQAKISLF 20  
 ||||| : |||||  
 Db 301 VLPFRKMKGLFSQAKISLF 320

RESULT 14  
 WMB57.  
 ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - human herpesvirus 1  
 N:Alternate names: ribonucleotide reductase small chain  
 C:Species: human herpesvirus 1  
 C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 16-Jun-2000  
 C:Accession: D30088  
 R:McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; P  
 J. Gen. Virol. 69, 1531-1574, 1988  
 A:Title: The complete DNA sequence of the long unique region in the genome of herpes  
 A:Reference number: A30083; MUID:88274327; PMID:2839594  
 A:Accession: D30088  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA



A:Residues: 1-340 <MCG>  
 A:Cross-references: GB:X14112; NID:g1944536; PIDN:CAA32303.1; PID:g59540; GB:D00317  
 C:Genetics:  
 A:Gene: ULA0  
 C:Function:  
 A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside  
 C:Superfamily: herpesvirus ribonucleoside-diphosphate reductase small chain  
 C:Keywords: deoxyribonucleotide biosynthesis; early protein; iron; metalloprotein; oxidc  
 F:94,124,127,187,221,224/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, Glu, His) #sta  
 F:131/Active site: Tyr (stable tyrosyl radical) #status predicted  
  
 Query Match 43.6%; Score 44; DB 1; Length 340;  
 Best Local Similarity 61.5%; Pred. NO. 16;  
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
 OY 3 PPRKMKGLFSOA 15  
 |||::|||  
 DB 242 PPDPRVYGLFRQA 254  
  
 ULA 15  
 S74937  
 hypothetical protein slr0688 - *Synechocystis* sp. (strain PCC 6803)  
 C:Species: *Synechocystis* sp.  
 A:Variety: PCC 6803  
 C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
 C:Accession: S74937  
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
 O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
 DNA Res. 3, 109-136, 1996  
 A>Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*  
 S.  
 A:Reference number: S74322; MUID:97061201; PMID:8905231  
 A:Accession: S74937  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-343 <KAN>  
 A:Cross-references: EMBL:D90902; GB:AB001339; NID:g1652027; PIDN:BA16977.1; PID:g101771  
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
  
 Query Match 43.6%; Score 44; DB 2; Length 343;  
 Best Local Similarity 47.1%; Pred. NO. 16;  
 Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;  
  
 OY 3 PPRKMKGLFSOA 19  
 |||::|||  
 DB 240 PPDPRCKGLFRGAETI 256  
  
 Search completed: October 1, 2003, 10:04:36  
 Job time : 15.8235 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 ; Search time 6.58824 seconds  
(without alignments)  
142.760 Million cell updates/sec

Title: US-09-171-432a-42

Perfect score: 101

Sequence: 1 VLPPRRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	852	1	P06442 hepatitis a
2	101	100.0	2227	1	P08617 hepatitis a
3	101	100.0	2227	1	P06441 hepatitis a
4	96	95.0	2226	1	P06441 hepatitis a
5	96	95.0	2226	1	P26580 hepatitis a
6	96	95.0	2226	1	P26581 hepatitis a
7	96	95.0	2230	1	P14553 simian hepa
8	95	94.1	2227	1	P13901 hepatitis a
9	84	83.2	839	1	P31788 simian hepa
10	75	73.3	341	1	P013672 hepatitis a
11	52.5	52.0	1025	1	Q12882 homo sapien
12	45.5	45.0	1025	1	Q28007 bos tauris
13	45.5	45.0	1025	1	Q28007 bos tauris
14	45.5	45.0	1025	1	Q28007 bos tauris
15	44.6	44.6	178	1	Q914d4 aradidopsis
16	44.6	44.6	340	1	P10224 herpes simp
17	44.6	44.6	340	1	P06474 herpes simp
18	43	42.6	389	1	P35515 methanosarc
19	43	42.6	587	1	O53609 streptomyc
20	42	41.6	178	1	P51418 aradidopsis
21	41	40.6	289	1	P43688 mus musculi
22	41	40.6	359	1	O54873 citiculus
23	41	40.6	1040	1	P22063 rattus norv
24	41	40.6	1234	1	O62315 mus musculi
25	40	39.6	135	1	P45777 haemophilus
26	40	39.6	292	1	P30647 caenorhabdit
27	40	39.6	310	1	P31330 mus musculi
28	40	39.6	312	1	Q12904 homo sapien
29	40	39.6	353	1	Q9h224 homo sapien
30	40	39.6	354	1	O9egm3 mus musculi
31	40	39.6	371	1	P43698 canis famli
32	40	39.6	372	1	P50220 homo sapien
33	40	39.6	372	1	P23441 rattus norv

34	40	39.6	448	1	DHE4_HELPY	P55990 helicobacte
35	40	39.6	523	1	YMR9_YEAST	Q05040 saccharomyc
36	40	39.6	620	1	ARR8_SCHPO	Q9us07 schizosacch
37	40	39.6	1048	1	SR44_RAT	O63627 rattus norv
38	40	39.6	1157	1	SR44_HUMAN	O95104 homo sapien
39	39.5	39.1	1039	1	M2C1_MOUSE	O95189 mus musculi
40	39.5	39.1	1040	1	M2C1_RAT	P21139 rattus norv
41	39.5	38.6	178	1	RLIX_CASSA	O9atf5 castanea sa
42	39	38.6	204	1	RS4_TREPA	O83328 treponema p
43	39	38.6	347	1	E13A_SOYBN	O03773 glycine max
44	39	38.6	427	1	TRB1_AERPE	O9y8t5 aeropyrum p
45	39	38.6	462	1	WDR8_MOUSE	O9jmr8 mus musculi

## ALIGNMENTS

RESULT 1	ID	POLG_HPVC	STANDARD:	PRT:	852 AA.
AC	P06442	083741; 083742;			
DT	01-JAN-1988	(rel. 06, Created)			
DT	01-JAN-1988	(rel. 06, Last sequence update)			
DT	16-OCT-2001	(rel. 40, Last annotation update)			
DE	Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).				
OS	Hepatitis A virus (strain CR326).				
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;				
OC	Hepatovirus.				
OX	NCBI_TaxID=12097;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=85185648; PubMed=2985793;				
RA	Linemeyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V.,				
RA	Young A., Mitra S.W.,				
RT	"Molecular cloning and partial sequencing of hepatitis A viral cDNA."				
RL	J. Virol. 54:247-255(1985).				
CC	-1 SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,				
CC	EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,				
CC	VP3, AND VP4.				
CC	-1 PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.				
CC	-1 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sdb.ch/announce/">http://www.isb-sdb.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sdb.ch">license@sdb.ch</a> ).				
CC	-----				
DR	EMBL: M10033; AAA5470.1; -				
DR	PIR: A03904; GNNYHA.				
KW	Polyprotein; Coat protein; Core protein.				
FT	CHAIN 1 23				
FT	CHAIN 24 245				
FT	CHAIN 246 491				
FT	CHAIN 492 836				
FT	CHAIN 837 >852				
FT	NON_TER 852 852				
FT	SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820E CRC64;				
Query Match	100.0%; Score 101; DB 1; Length 852;				
Best Local Similarity	100.0%; Pred. No. 9.9e-09;				
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Oy	1 VLPPRRKMKGLFSQAKISLF 20				
Db	823 VLPPRRKMKGLFSQAKISLF 842				
RESULT 2					
POLG_HPVC					

ID POLG\_HPAVL STANDARD; PRT; 2227 AA.  
AC P06617; P06643; Q81082;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Genome polypeptide [containing: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
DE Hepatitis A virus (strain HM-175).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.  
OC NCBI\_TaxID=12098;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Wild type;  
RX MEDLINE=87061253; PubMed=3023706;  
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A., Baroudy B.M.;  
RT "Complete nucleotide sequence of wild-type hepatitis A virus: comparison with different strains of hepatitis A virus and other picornaviruses.";  
RT J. Virol. 61:50-59(1987).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Attenuated;  
RX MEDLINE=87175701; PubMed=3031686;  
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstone S.M., Purcell R.H.;  
RT "Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with wild-type virus.";  
RT Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).  
[3]  
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.  
RX MEDLINE=85166289; PubMed=2984684;  
RA Baroudy B.M., Ticehurst J.R., Mele T.A., Matzel J.V. Jr., Purcell R.H., Feinstone S.M.;  
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA polymerase.";  
RT Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).  
RL -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate + (RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS. MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT SHOWN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M14114; AAA45475.1; -;  
DR EMBL; M14707; AAA45465.1; -;  
DR EMBL; M14707; AAA45466.1; ALT-INT.  
DR EMBL; M16632; AAA45471.1; -;  
DR PIR; A25981; GNNYHM.  
DR PIR; A94149; GNNYHM.  
DR PDB; 1HAV; 23-DEC-96.  
DR MEROPS; C03.005; -;  
DR InterPro; IPR004004; Calic1\_pol\_hel.  
DR InterPro; IPR000605; RNA\_helicase.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR InterPro; IPR007094; RNA\_pol\_PSV1r.  
DR Pfam; PF00660; RNA\_dep\_RNA\_pol; 1.

DR Pfam; PF00910; RNA\_helicase; 1.  
DR PRINTS; PR00918; CALICVIRUSNS.  
KM Polypeptide; Coat protein; Core protein; Transferase; Hydrolase; Ribol protease; 3D-structure.  
DE RNA-directed RNA polymerase;  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 836  
FT CHAIN 837 980  
FT CHAIN 981 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1496  
FT CHAIN 1497 1519  
FT CHAIN 1520 1738  
FT CHAIN 1739 2227  
FT VARIANT 764 764  
FT VARIANT 821 821  
FT VARIANT 1052 1052  
FT VARIANT 1062 1062  
FT VARIANT 1118 1118  
FT VARIANT 1151 1151  
FT VARIANT 1163 1163  
FT VARIANT 1277 1277  
FT VARIANT 1500 1500  
FT VARIANT 1805 1805  
FT VARIANT 1930 1930  
SQ SEQUENCE .2227 AA; 251506 MW; 01E225E7AEB740A6 CRC64;  
Query Match 100.0%; Score 101; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.7e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMKGLFSQAKISLF 20  
|||||  
Db 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 3  
POLG\_HPAVL STANDARD; PRT; 2227 AA.  
AC P06641;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polypeptide [containing: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
DE Hepatitis A virus (strain LA).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.  
OC NCBI\_TaxID=12099;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85190549; PubMed=2986127;  
RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A., Merryweather J., van Nest G., Dina D.;  
RT "Primary structure and gene organization of human hepatitis A virus.";  
RT Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).  
RL -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate + (RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS. MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

```

CC or send an email to license@sb-stb.ch).
CC -----
DR EMBL: K02990; AAA5472.1; -.
DR PIR: A03903; GNNYR.
DR MEROPS: C03.005; -.
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINTS: PR00918; CALICVIRUSNS.
DR PolyProtein: Coat protein; Core protein; Transferase;
KM RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).
FT CHAIN 837 980 CORE PROTEIN P2A.
FT CHAIN 981 1072 CORE PROTEIN P2B.
FT CHAIN 1077 1422 CORE PROTEIN P2C.
FT CHAIN 1423 1484 PROBABLE PROTEIN P3A.
FT CHAIN 1485 1507 PROBABLE PROTEIN P3B.
FT CHAIN 1508 1678 PROBABLE PROTEIN P3C.
FT CHAIN 1679 2227 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.7e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 4
POLG_HPAV2 STANDARD; PRT; 2226 AA.
ID POLG_HPAV2
AC P26580;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 24a).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12094;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE-91162758; PubMed-1705995;
RX Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromean S.T., Jansen R.W.;
RT Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.
RT J. Virol. 65:2056-2065(1991).
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (see http://www.isb-stb.ch/announce/
CC or send an email to license@sb-stb.ch).
CC -----
DR EMBL: M59810; AAA5468.1; -.
DR MEROPS: C03.005; -.
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINTS: PR00918; CALICVIRUSNS.
DR PolyProtein: Coat protein; Core protein; Transferase;
KM RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).
FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).
FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).
FT CHAIN 492 794 COAT PROTEIN VP1 (P1D).
FT CHAIN 795 900 CORE PROTEIN P2A.
FT CHAIN 901 1087 CORE PROTEIN P2B.
FT CHAIN 1088 1432 CORE PROTEIN P2C.
FT CHAIN 1433 1495 PROBABLE PROTEIN P3A.
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

Query Match 95.0%; Score 96; DB 1; Length 2226;
Best Local Similarity 95.0%; Pred. No. 1.8e-07;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLPPPRKMKGLFSQAKISLF 20
DB 823 VLPPPRKMKGLFSQAKISLF 842

RESULT 5
POLG_HPAV4 STANDARD; PRT; 2226 AA.
ID POLG_HPAV4
AC P26581;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 43c).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12095;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE-91162758; PubMed-1705995;
RX Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromean S.T., Jansen R.W.;
RT Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination.
RT J. Virol. 65:2056-2065(1991).
RL J. Virol. 65:2056-2065(1991).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M59809; AAA45469.1; -  
DR MEROPS; C03.005; -  
DR InterPro; IPR000605; RNA\_helicase.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR InterPro; IPR007094; RNA\_pol\_PSVL.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
KW Polypeptide; Coat protein; Core protein; Transferase;  
RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 794  
FT CHAIN 795 900  
FT CHAIN 901 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1495  
FT CHAIN 1496 1518  
FT CHAIN 1519 1737  
FT CHAIN 1738 2226  
SQ SEQUENCE 2226 AA; 251107 MW; 403B4C80B09BF75 CRC64;

Query Match 95.0%; Score 96; DB 1; Length 2226;  
Best Local Similarity 95.0%; Pred. No. 1.8e-07;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLPPRRMKGLFSQAKISLF 20  
DB 823 VLPPRRMKGLFSQAKISLF 842

RESULT 6  
POLG\_HPAV8 STANDARD; PRT: 2226 AA.  
AC P26582;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins  
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
P3D (EC 2.7.7.48)].  
DE P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain 18f).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
NCBI\_Taxid=12096;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=91162756; PubMed=1705995;  
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,  
RA Crommons T., Jansen R.W.;  
RT "Antigenic and genetic variation in cyclopathic hepatitis A virus  
RT variants arising during persistent infection: evidence for genetic  
RT recombination."  
RT J. Virol. 65:2056-2065(1991).  
RL J. Virol. 65:2056-2065(1991).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC (RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; M59808; AAA45467.1; -  
DR PDB; 1QAT; 15-MAY-00.  
DR MEROPS; C03.005; -  
DR InterPro; IPR000605; RNA\_helicase.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR InterPro; IPR007094; RNA\_pol\_PSVL.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
KW Polypeptide; Coat protein; Core protein; Transferase;  
RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 794  
FT CHAIN 795 900  
FT CHAIN 901 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1495  
FT CHAIN 1496 1518  
FT CHAIN 1519 1737  
FT CHAIN 1738 2226  
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D68 CRC64;

Query Match 95.0%; Score 96; DB 1; Length 2226;  
Best Local Similarity 95.0%; Pred. No. 1.8e-07;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLPPRRMKGLFSQAKISLF 20  
DB 823 VLPPRRMKGLFSQAKISLF 842

RESULT 7  
POLG\_HPAV5 STANDARD; PRT: 2230 AA.  
AC P14533;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins  
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
P3D (EC 2.7.7.48)].  
DE P3D (EC 2.7.7.48)].  
OS Simian hepatitis A virus (strain AGM-27).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
NCBI\_Taxid=12102;  
[1]  
SEQUENCE FROM N.A.  
MEDLINE=91311420; PubMed=1649901;  
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,  
RA Purcell R.H.;  
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome  
RT structure and growth in cell culture with other HAV strains."  
RT J. Gen. Virol. 72:1677-1683(1991).  
RL J. Gen. Virol. 72:1677-1683(1991).  
RN [2]  
RP SEQUENCE OF 1750-2164 FROM N.A.  
RX MEDLINE=89232168; PubMed=2541023;  
RA Balayan M.S., Kusov Y.Y., Andapapartz A.G., Tsarev S.A.,  
RA Sverdlov E.D., Chizhikov V.E., Bilinov V.M., Vasilenko S.K.;  
RT "Variations in genome fragments coding for RNA polymerase in human  
RT and simian hepatitis A viruses."  
RT FEBS Lett. 247:425-428(1989).  
RL FEBS Lett. 247:425-428(1989).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC (RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: D00924; BAA00766.1; -  
DR EMBL: X15461; CAA33490.1; -  
DR PIR: A30470; GNRYS.A.  
DR MEROPS: C03.005; -  
DR InterPro: IPR004004; Calic\_pol\_hel.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR007094; RNA\_pol\_PSVir.  
DR Pfam: PF00680; RNA\_dep\_RNA\_pol.1.  
DR Pfam: PF00910; RNA\_helicase.1.  
DR PRINTS: PR00918; CALICVIRUSNS.  
DR Polyprotein: Coat protein; Core protein; Transferase;  
DR RNA-directed RNA polymerase; Hydrolyase; Thiol protease.  
FT CHAIN 1 27  
FT CHAIN 28 249 COAT PROTEIN VP4 (P1A).  
FT CHAIN 250 495 COAT PROTEIN VP2 (P1B).  
FT CHAIN 496 795 COAT PROTEIN VP3 (P1C).  
FT CHAIN 796 984 COAT PROTEIN VP1 (P1D).  
FT CHAIN 985 1091 CORE PROTEIN P2A.  
FT CHAIN 1092 1426 CORE PROTEIN P2B.  
FT CHAIN 1427 1498 CORE PROTEIN P2C.  
FT CHAIN 1499 1521 PROBABLE PROTEIN P3A.  
FT CHAIN 1522 1741 PROBABLE PROTEIN P3B.  
FT CHAIN 1742 2230 PROBABLE PROTEIN P3C.  
FT CHAIN 2230 251296 RNA-DIRECTED POLYMERASE 3D.  
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 95.0%; Score 96; DB 1; Length 2230;  
Best Local Similarity 90.0%; Pred. No. 1.8e-07;  
Matches 18; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRKKKGLFSQAKISLF 20  
DB 827 ILPPRRKKKGLFSQAKISLF 846

RESULT 8  
POLG\_HPAVM STANDARD; PRT; 2227 AA.  
AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;  
AC Q81090; Q81091; Q81092; Q81093;  
01-JAN-1990 (Rel. 13; Created)  
01-JAN-1990 (Rel. 13; Last sequence update)  
28-FEB-2003 (Rel. 41; Last annotation update)  
DE Genome polyprotein (Contains: Coat proteins VP1 TO VP4; Core proteins  
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
DE P3D (EC 2.7.7.48)).  
OS Hepatitis A virus (strain MBB).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12100;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86045071; PubMed=2823500;  
RA Paul A.V., Tada H., der Helm K., Wiesel T., Klein R., Wimmer E.,  
RA Deinhardt F.;  
RT "The entire nucleotide sequence of the genome of human hepatitis A  
RT virus (isolate MBB).";  
RL Virus Res. 8:153-171(1987).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC {RNA}(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMTIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT

CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: M20273; AAA45474.1; -  
DR MEROPS: C03.005; -  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR007094; RNA\_pol\_PSVir.  
DR Pfam: PF00680; RNA\_dep\_RNA\_pol.1.  
DR Pfam: PF00910; RNA\_helicase.1.  
DR Polyprotein: Coat protein; Core protein; Transferase;  
DR RNA-directed RNA polymerase; Hydrolyase; Thiol protease.  
KW CHAIN 1 23  
KW CHAIN 24 245 COAT PROTEIN VP4 (P1A).  
KW CHAIN 246 491 COAT PROTEIN VP2 (P1B).  
KW CHAIN 492 836 COAT PROTEIN VP3 (P1C).  
KW CHAIN 837 980 COAT PROTEIN VP1 (P1D).  
KW CHAIN 981 1087 CORE PROTEIN P2A.  
KW CHAIN 1088 1422 CORE PROTEIN P2B.  
KW CHAIN 1423 1496 CORE PROTEIN P2C.  
KW CHAIN 1497 1519 PROBABLE PROTEIN P3A.  
KW CHAIN 1520 1738 PROBABLE PROTEIN P3B.  
KW CHAIN 1739 2227 PROBABLE PROTEIN P3C.  
KW CHAIN 2227 251425 RNA-DIRECTED POLYMERASE 3D.  
SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match 94.1%; Score 95; DB 1; Length 2227;  
Best Local Similarity 95.0%; Pred. No. 2.6e-07;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLPPRRKKKGLFSQAKISLF 20  
DB 823 VLPPRRKKKGLFSQAKISLF 842

RESULT 9  
POLG\_HPAVT STANDARD; PRT; 839 AA.  
AC P31788;  
DT 01-JUL-1993 (Rel. 26; Created)  
DT 01-JUL-1993 (Rel. 26; Last sequence update)  
DT 16-OCT-2001 (Rel. 40; Last annotation update)  
DE Genome polyprotein (Contains: Coat proteins VP1 TO VP4; Core protein  
DE P2A) (Fragment).  
OS Simian hepatitis A virus (strain CY-145).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=31707;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91311421; PubMed=1649902;  
RA Nalnan O.V., Margolis H.S., Robertson B.H., Balayan M., Brinton M.A.;  
RA "Sequence analysis of a new hepatitis A virus naturally infecting  
RT cynomolgus macaques (Macaca fascicularis).";  
RL J. Gen. Virol. 72:1685-1689(1991).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMTIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M59286; AAA45473.1; -  
CC PIR: J01180; GNTY52.  
-----

KM Polypeptide: Coar protein; Core protein.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 2  
FT CHAIN ? >839  
FT NON\_TER 839 839  
SQ SEQUENCE 839 AA; 93825 MW; 2CACCB1E192DEC CRC64;

Query Match 83.2%; Score 84; DB 1; Length 839;  
Best Local Similarity 83.3%; Pred. No. 6.1e-06;  
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMGIFSOAKIS 18  
:|||||:|||||:  
822 ILPPPRKMGIFSOAKIS 839

RESULT 10  
POLG\_HPAV1 STANDARD; PRT; 341 AA.  
AC P13672;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Genome polypeptide [Contains: Coar proteins VP1 TO VP3; Core protein  
DE P2A] (Fragment).  
OS Hepatitis A virus (strain LCD-1).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OX NCBI\_Taxid=12093;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-89263805; PubMed-2542903;  
RA Andonov A.P., Lau P., Chaudhary R.;  
RT "Nucleotide sequence of the VP1 gene from a Chinese strain of  
RT hepatitis A virus (HAV)."  
RL Nucleic Acids Res. 17:3594-3594(1989).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X14666; CA32794.1; -  
CC PIR: S04137; S04137.  
KM Polypeptide: Coar protein; Core protein.  
FT CHAIN 1 1  
FT CHAIN 2 340  
FT CHAIN 341 >341  
FT CHAIN 341 341  
FT NON\_TER 341 341  
SQ SEQUENCE 341 AA; 38003 MW; 066918289BF126D5 CRC64;

Query Match 74.3%; Score 75; DB 1; Length 341;  
Best Local Similarity 93.3%; Pred. No. 7.1e-05;  
Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKMGIFSOA 15  
:|||||:|||||:  
327 VLPPPRKMGIFSOA 341

RESULT 11  
DPYD\_HUMAN STANDARD; PRT; 1025 AA.  
ID DPYD\_HUMAN  
AC Q12882; Q16694; Q16761;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Dihydropyrimidine dehydrogenase [NADP+] precursor (EC 1.3.1.2) (DPD)  
DE (dihydase) (dihydrouacil dehydrogenase) (dihydrochymine  
DE dehydrogenase).  
GN DPYD.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Liver;  
RX MEDLINE-94365020; PubMed-8083224;  
RA Yokota H., Fernandez-Salguero P., Furuya H., Lin K., McBride O.W.,  
RA Podeschun B., Schnackerz K.D., Gonzalez F.J.;  
RT "cDNA cloning and chromosome mapping of human dihydropyrimidine  
RT dehydrogenase, an enzyme associated with 5-fluorouracil toxicity and  
RT congenital thymine uraciluria."  
RL J. Biol. Chem. 269:23192-23196(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-97280676; PubMed-9135003;  
RA Johnson M.R., Wang K., Tillmanns S., Albin N., Diasio R.B.;  
RT "Structural organization of the human dihydropyrimidine dehydrogenase  
RT gene."  
RL Cancer Res. 57:1660-1663(1997).  
RN [3]  
RP SEQUENCE OF 581-635 FROM N.A.  
RC TISSUE-Liver;  
RX MEDLINE-97047101; PubMed-8892022;  
RA Vreken P., van Kullenburg A.B.P., Meisema R., Smit G.P.A.,  
RA Bakker H.D., de Abreu R.A., van Gennip A.H.;  
RT "A point mutation in an invariant splice donor site leads to exon  
RT skipping in two unrelated Dutch patients with dihydropyrimidine  
RT dehydrogenase deficiency."  
RL J. Inherit. Metab. Dis. 19:645-654(1996).  
RN [4]  
RP CHARACTERIZATION, AND PARTIAL SEQUENCE.  
RC TISSUE-Liver;  
RX MEDLINE-92381021; PubMed-1512248;  
RA Lu Z.-H., Zhang R., Diasio R.B.;  
RT "Purification and characterization of dihydropyrimidine dehydrogenase  
RT from human liver."  
RL J. Biol. Chem. 267:17102-17109(1992).  
RN [5]  
RP VARIANTS ARG-29; TRP-235 AND HIS-886.  
RX MEDLINE-98102836; PubMed-9439663;  
RA Vreken P., van Kullenburg A.B.P., Meisema R., van Gennip A.H.;  
RT "Dihydropyrimidine dehydrogenase (DPD) deficiency: identification and  
RT expression of missense mutations C29R, R886H and R235W."  
RL Hum. Genet. 101:333-338(1997).  
RN [6]  
RP VARIANTS ARG-29; TRP-235 AND HIS-886.  
RX MEDLINE-97411371; PubMed-9266349;  
RA Vreken P., van Kullenburg A.B.P., Meisema R., van Gennip A.H.;  
RT "Identification of novel point mutations in the dihydropyrimidine  
RT dehydrogenase gene."  
RL J. Inherit. Metab. Dis. 20:335-338(1997).  
CC -1- FUNCTION: INVOLVED IN PYRIMIDINE BASE DEGRADATION. CATALYZES THE  
CC REDUCTION OF URACIL AND THYMINE. ALSO INVOLVED THE DEGRADATION OF  
CC THE CHEMOTHERAPEUTIC DRUG 5-FLUOROURACIL.  
CC -1- CATALYTIC ACTIVITY: 5,6-dihydrouacil + NADP(+) = uracil + NADPH.  
CC -1- COFACTOR: TWO EACH OF FAD AND FMN. ALSO CONTAINS TWO 4FE-4S  
CC CLUSTERS. CONTAINS APPROXIMATELY 33 IRON ATOMS PER MOLECULE.  
CC -1- PATHWAY: INITIAL AND RATE-LIMITING ENZYME IN THE 3-STEP PATHWAY OF



```
FT FT VARIANT 543 543 I -> V (IN ALLELE DYPD+5; dbsNP:1801159).  
FT FT VARIANT 732 732 /FtId-VAR_005176  
V->I (IN dBSNP:1801160).  
FT FT VARIANT 886 886 R -> H (IN ALLELE DPYD+9B; 25% OF  
ACTIVITY; dbsNP:1801267) .  
FT FT VARIANT 995 995 V -> F (IN ALLELE DPYD+10; LOW ACTIVITY;  
dbsNP:1801268)  
FT FT CONFLICT 910 910 /FtId-VAR_005178.  
S -> N (IN REF. 2).  
SO SEQUENCE 1025 AA; 111374 MM; 52IC9430C7CF69AFA CRC64;  
  
Query Match Best Local Similarity 52.0%; Score 52.5; DB 1; Length 1025;  
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;  
  
Qy 2 LPPRKKMKGLFSQAISLIF 20  
||| || | : ||| :: ||  
Db 176 LPPEKMSSEAVS-AKIALF 193  
  
RESULT 12  
DPYD_BOVIN STANDARD; PRT; 1025 AA.  
ID ID DPYD_BOVIN STANDARD; PRT; 1025 AA.  
AC Q28007; O9TRV4;  
DT 01-NOV-1997 (Rel. 35, Created)  
DI 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Dihydroxyrimidine dehydrogenase [NADP+] (EC 1.3.1.2) (DDP)  
DE (DHBPdhase) (Dihydrouacil dehydrogenase) (Dihydrothyamine  
dehydrogenase) .  
DE DPYD.  
GN Bos taurus (Bowling).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
CC Bovidae; Bovinae; Bos.  
OX NCBI_Taxid=9913;  
RN [1]  
RM SEQUENCE FROM N.A.  
RP TISSUE-Layer:  
RX MEDLINE-97069948; PubMed-8912928;  
RA Aldin N., Johnson M.R., Diasio R.B.;  
RT "cDNA cloning of bovine liver dihydropyrimidine dehydrogenase.*";  
RL DNA Seq. 6:243-250(1996).  
[2]  
RS SEQUENCE OF 668-678, AND ENZYME REGULATION.  
RC TISSUE-layer:  
RE MEDLINE-97041818; Pubmed-1939061.  
RX Porter D.J.T., Chestnut W.G., Taylor L.C.E., Merrill B.M., Spector T.;  
RA "Inactivation of dihydropyrimidine dehydrogenase by 5-Iodouracil." ;  
RT J. Biol. Chem. 266:19988-19994(1991).  
CC -1 FUNCTION: INVOLVED IN PYRIMIDINE BASE DEGRADATION. CATALYZES THE  
REDUCTION OF URACIL AND THYMINE.  
CC -1 CATABOLIC ACTIVITY: 5,6-dihydrouracil + NADP(+)= uracil + NADPH.  
CC -1 COPROCTOR: FAD AND FMN. ALSO CONSTRAINT TWO AFE-4S CLUSTERS.  
CC -1 ENZYME REGULATION: INACTIVATED BY 5-IDOURACIL.  
CC CC PATHWAY: INITIAL AND RATE-LIMITING ENZYME IN THE 3-STEP PATHWAY OF  
URACIL AND THYMINE CATABOLISM AND IN THE PATHWAY LEADING TO THE  
FORMATION OF BETA-ALANINE.  
CC -1 SUBUNIT: Homodimer (by similarity).  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
CC at the European Bioinformatics Institute. There are no restrictions on ways  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial/  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to licenses@isb.slb.ch).  
-----  
OR EMBL; UZ0981; ABAB40985.1; .  
OR InterPro; IPR0001450; 4Fe4S_ferredoxin.  
OR InterPro; IPR000759; Adenyl_reductase.
```

DR InterPro: IPR001295; DHO\_dh.  
DR InterPro: IPR005720; DHO\_dhl.  
DR InterPro: IPR001327; FAD\_pyr\_redox.  
DR InterPro: IPR003009; FMN\_enzyme.  
DR InterPro: IPR000103; Pyridine\_redox\_2.  
DR Pfam: PF01180; DHOdehase; 1.  
DR Pfam: PF00037; fer4; 2.  
DR Pfam: PF00070; Pyr\_redox; 1.  
DR PRINTS: PR00368; FADPNR.  
DR PRINTS: PR00469; PNDROTASE1.  
DR TIGRFAMs: TIGR01037; PYRD\_subl\_fam; 1.  
DR PROSITE: PS00196; 4FE4S\_FERRDOXIN; 1.  
DR PROSITE: PS00196; 4FE4S\_FERRDOXIN; 1.  
KW Oxidoreductase; NADP; Flavoprotein; FAD; FMN; Iron-sulfur; 4Fe-4S.  
FT METAL 953 953 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  
FT METAL 956 956 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  
FT METAL 959 959 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  
FT METAL 963 963 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  
FT METAL 986 986 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).  
FT METAL 989 989 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).  
FT METAL 992 992 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).  
FT METAL 996 996 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).  
SEQUENCE 1025 AA; 11696 MW; 5855F93A06C47E4F CRC64;

Query Match 45.0%; Score 45.5; DB 1; Length 1025;  
Best Local Similarity 61.1%; Pred. No. 16;  
Matches 11; Conservative 2; Mismatches 4; Indels 1; Gaps 1;  
2 LPPRRKMGKLFSSQAKISL 19  
Db 176 LPPRRKMGKLFSSQAKISL 192

RESULT 13  
RL1X\_ORYSA STANDARD; PRT; 178 AA.  
AC 0943F3;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Ribosomal protein L18a.  
GN RPL18A.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Erihartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
SEQUENCE FROM N.A.  
STRAIN=cv. Nipponbare;  
Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
clone:P0046E05";  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AP003237; BAB67920.1; -  
DR EMBL: AP003249; BAB89536.1; -  
DR Gramene: O943F3; -  
DR InterPro: IPR002670; Ribosomal\_L18ae.  
DR Pfam: PF01775; Ribosomal\_L18ae; 1.  
DR Ribosomal protein.  
KW SEQUENCE 178 AA; 21400 MW; 1A508A4670C8D93B CRC64;

Query Match 44.6%; Score 45; DB 1; Length 178;  
Best Local Similarity 45.0%; Pred. No. 3;

Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
QY 1 VLPPRRKMGKLFSSQAKISLF 20  
Db 158 VRPPTRKTKTFKSRPNLF 177

RESULT 14  
RL1Y\_ARATH STANDARD; PRT; 178 AA.  
AC OryL18A;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 60S ribosomal protein L18a-2.  
GN RPL18AB OR AR3G14600 OR MTE1.10.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
SEQUENCE FROM N.A.  
STRAIN=cv. Columbia;  
RC MEDLINE-20277480; PubMed-10819329;  
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence  
RT features of the regions of 4,504,864 bp covered by sixty pl and TAC  
RT clones";  
RL DNA Res. 7:131-135(2000).  
RN [2]  
SEQUENCE FROM N.A.  
STRAIN=cv. Columbia;  
RC Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;  
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the  
RT SSP consortium (Salk/Stanford/RGC)";  
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE L18AE FAMILY OF RIBOSOMAL PROTEINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AB023038; BAB02392.1; -  
DR EMBL: AY072340; AAL60048.1; -  
DR InterPro: IPR002670; Ribosomal\_L18ae.  
DR Pfam: PF01775; Ribosomal\_L18ae; 1.  
KW Ribosomal protein.  
SO SEQUENCE 178 AA; 21309 MW; 28A1B900F3C4402B CRC64;

Query Match 44.6%; Score 45; DB 1; Length 178;  
Best Local Similarity 45.0%; Pred. No. 3;  
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 VLPPRRKMGKLFSSQAKISLF 20  
Db 158 VRPPTRKTKTFKSRPNLF 177

RESULT 15  
RLR2\_HSV11 STANDARD; PRT; 340 AA.  
AC P10224;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 01-AUG-1990 (Rel. 15, Last annotation update)  
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)  
DE (Ribonucleotide reductase) (38 kDa subunit).  
GN U1A0.

```

OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_Taxid-10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88274327; PubMed-2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT *The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1.*;
RL J. Gen. Virol. 69:1531-1574(1988).
CC -1- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thioresoxin + H(2)O -> ribonucleoside diphosphate + reduced
CC thioresoxin.
CC -1- COFACTOR: BINDS 2 IRON IONS.
CC -1- PATHWAY: DNA replication pathway; first step.
CC -1- SUBUNIT: Heterodimer of a large and a small chain.
CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14112; CAA32303.1; -.
DR PIR: D30088; WMBE57.
DR InterPro: IPR000358; Ribonuc_redctse.
DR Pfam: PF00268; ribonuc_red_sm; 1.
DR PROSITE: PS00368; RIBORED_SMALL; 1.
KW Oxidoreductase; DNA replication; Iron.
FT METAL 94 94 IRON 1 (BY SIMILARITY).
FT METAL 124 124 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 127 127 IRON 1 (BY SIMILARITY).
FT METAL 187 187 IRON 2 (BY SIMILARITY).
FT METAL 221 221 IRON 2 (BY SIMILARITY).
FT METAL 224 224 IRON 2 (BY SIMILARITY).
FT ACT_SITE 131 131 BY SIMILARITY.
SQ SEQUENCE 340 AA; 38019 MW; 4B4ED994BF74FD3F CRC64;

```

```

Query Match 43.6%; Score 44; DB 1; Length 340;
Best Local Similarity 61.5%; Pred. No. 8.7;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
3 PPRKMKGLFSQA 15
||||:|||||
242 PPDNRYGLFRQA 254

```

Search completed: October 1, 2003, 09:57:46  
Job time : 7.58824 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 33.1765 Seconds  
(without alignments)  
155.564 Million cell updates/sec

Title: US-09-171-432a-42  
Perfect score: 101  
Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
al number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_23:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	94	12	Q9ENT9 hepatitis a
2	101	100.0	94	12	Q9ENT9 hepatitis a
3	101	100.0	94	12	Q9ENT9 hepatitis a
4	101	100.0	94	12	Q9ENT9 hepatitis a
5	101	100.0	94	12	Q9ENT9 hepatitis a
6	101	100.0	94	12	Q9ENT9 hepatitis a
7	101	100.0	94	12	Q9ENT9 hepatitis a
8	101	100.0	94	12	Q9ENT9 hepatitis a
9	101	100.0	94	12	Q9ENT9 hepatitis a
10	101	100.0	94	12	Q9ENT9 hepatitis a
11	101	100.0	94	12	Q9ENT9 hepatitis a
12	101	100.0	94	12	Q9ENT9 hepatitis a
13	101	100.0	94	12	Q9ENT9 hepatitis a
14	101	100.0	94	12	Q9ENT9 hepatitis a
15	101	100.0	94	12	Q9ENT9 hepatitis a
16	101	100.0	94	12	Q9ENT9 hepatitis a

17	101	100.0	94	12	Q9ENT9 hepatitis a
18	101	100.0	94	12	Q9ENT9 hepatitis a
19	101	100.0	94	12	Q9ENT9 hepatitis a
20	101	100.0	94	12	Q9ENT9 hepatitis a
21	101	100.0	94	12	Q9ENT9 hepatitis a
22	101	100.0	94	12	Q9ENT9 hepatitis a
23	101	100.0	115	12	Q9DMR4 hepatitis a
24	101	100.0	116	12	Q92941 hepatitis a
25	101	100.0	116	12	Q9W7X7 hepatitis a
26	101	100.0	116	12	Q9W7X7 hepatitis a
27	101	100.0	116	12	Q9W7X7 hepatitis a
28	101	100.0	116	12	Q9W7X7 hepatitis a
29	101	100.0	116	12	Q9W7X7 hepatitis a
30	101	100.0	116	12	Q9W7X7 hepatitis a
31	101	100.0	116	12	Q9W7X7 hepatitis a
32	101	100.0	126	12	Q9W7X7 hepatitis a
33	101	100.0	132	12	Q9W7X7 hepatitis a
34	101	100.0	132	12	Q9W7X7 hepatitis a
35	101	100.0	132	12	Q9W7X7 hepatitis a
36	101	100.0	132	12	Q9W7X7 hepatitis a
37	101	100.0	132	12	Q9W7X7 hepatitis a
38	101	100.0	132	12	Q9W7X7 hepatitis a
39	101	100.0	132	12	Q9W7X7 hepatitis a
40	101	100.0	132	12	Q9W7X7 hepatitis a
41	101	100.0	132	12	Q9W7X7 hepatitis a
42	101	100.0	132	12	Q9W7X7 hepatitis a
43	101	100.0	132	12	Q9W7X7 hepatitis a
44	101	100.0	132	12	Q9W7X7 hepatitis a
45	101	100.0	132	12	Q9W7X7 hepatitis a

## ALIGNMENTS

RESULT 1  
Q9ENT9 PRELIMINARY; PRT; 94 AA.  
AC Q9ENT9;  
DT 01-MAR-2001 (TREMBLrel, 16, Created)  
DT 01-MAR-2001 (TREMBLrel, 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel, 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Nagasaki 32;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases."  
DR EMBL; AB038298; BAB1836.1; -;  
FT NON\_TER 1 94  
FT NON\_TER 1 94  
SQ SEQUENCE 94 AA; 10807 MW; F19ACE91B8C4AFD CRC64;  
Query Match 100.0%; Score 101; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VLPPPRKMKGLFSQAKISLF 20  
DB 74 VLPPPRKMKGLFSQAKISLF 93  
RESULT 2  
Q9ENT9 PRELIMINARY; PRT; 94 AA.  
AC Q9ENT9;  
DT 01-MAR-2001 (TREMBLrel, 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, last sequence update)  
DE 01-MAR-2001 (Tremblrel. 16, last annotation update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;  
OC Hepatovirus.  
NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Nagasaki 03;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
infection: Prolonged HAV viremia and mild liver injury."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB038278; BAB1816.1; -  
FT NON\_TER 1 1  
FT SEQUENCE 94 AA; 10862 MW; 9AF9FDDAE8C4BAC CRC64;  
SQ

Query Match 100.0%; Score 101; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRKMGLEFSQAKISLF 20  
DB 74 VLPPRRKMGLEFSQAKISLF 93

RESULT 3  
Q9ENV5 PRELIMINARY; PRT; 94 AA.

AC 09ENV5;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;  
OC Hepatovirus.  
NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Nagasaki 08;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
infection: Prolonged HAV viremia and mild liver injury."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB038282; BAB1820.1; -  
FT NON\_TER 1 1  
FT SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;  
SQ

Query Match 100.0%; Score 101; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRKMGLEFSQAKISLF 20  
DB 74 VLPPRRKMGLEFSQAKISLF 93

RESULT 4  
Q9ENV2 PRELIMINARY; PRT; 94 AA.

AC 09ENV2;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;  
OC

OC Hepatovirus.  
NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Nagasaki 28;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
infection: Prolonged HAV viremia and mild liver injury."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB038295; BAB1833.1; -  
FT NON\_TER 1 1  
FT SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;  
SQ

Query Match 100.0%; Score 101; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRKMGLEFSQAKISLF 20  
DB 74 VLPPRRKMGLEFSQAKISLF 93

RESULT 5  
Q9ENV1 PRELIMINARY; PRT; 94 AA.

AC 09ENV1;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;  
OC Hepatovirus.  
NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Kantou 49;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
infection: Prolonged HAV viremia and mild liver injury."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB038306; BAB1844.1; -  
FT NON\_TER 1 1  
FT SEQUENCE 94 AA; 10842 MW; 741850915B86E1A6 CRC64;  
SQ

Query Match 100.0%; Score 101; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRKMGLEFSQAKISLF 20  
DB 74 VLPPRRKMGLEFSQAKISLF 93

RESULT 6  
Q9ENV3 PRELIMINARY; PRT; 94 AA.

AC 09ENV3;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses: ssRNA positive-strand viruses, no DNA stage: Picornaviridae;  
OC Hepatovirus.  
NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Nagasaki 12n;

RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
 RA Yasuoka A., Oka S.;  
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
 RT infection: Prolonged HAV viremia and mild liver injury.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB038284; BAB11822.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 94  
 SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKKMGLFSQAKISLF 20  
 DB 74 VLPPPRKKMGLFSQAKISLF 93

## RESULT 7

AC O9ENV1 PRELIMINARY; PRT; 94 AA.  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE Polypeptide (Fragment).  
 OS Hepatitis A virus.  
 OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 CC Hepatovirus.  
 RN NCBI\_TaxID=12092;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Nagasaki 29;  
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
 RA Yasuoka A., Oka S.;  
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
 RT infection: Prolonged HAV viremia and mild liver injury.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB038285; BAB11834.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 94  
 SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKKMGLFSQAKISLF 20  
 DB 74 VLPPPRKKMGLFSQAKISLF 93

## RESULT 8

AC O9ENV8 PRELIMINARY; PRT; 94 AA.  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE Polypeptide (Fragment).  
 OS Hepatitis A virus.  
 OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 CC Hepatovirus.  
 RN NCBI\_TaxID=12092;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Nagasaki 22;  
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
 RA Yasuoka A., Oka S.;  
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
 RT infection: Prolonged HAV viremia and mild liver injury.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB038289; BAB11827.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 94  
 SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKKMGLFSQAKISLF 20  
 DB 74 VLPPPRKKMGLFSQAKISLF 93

## RESULT 9

AC O9ENV2 PRELIMINARY; PRT; 94 AA.  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE Polypeptide (Fragment).  
 OS Hepatitis A virus.  
 OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 CC Hepatovirus.  
 RN NCBI\_TaxID=12092;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Nagasaki 15;  
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
 RA Yasuoka A., Oka S.;  
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
 RT infection: Prolonged HAV viremia and mild liver injury.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB038285; BAB11823.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 94  
 SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLPPPRKKMGLFSQAKISLF 20  
 DB 74 VLPPPRKKMGLFSQAKISLF 93

## RESULT 10

AC O9ENV5 PRELIMINARY; PRT; 94 AA.  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE Polypeptide (Fragment).  
 OS Hepatitis A virus.  
 OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 CC Hepatovirus.  
 RN NCBI\_TaxID=12092;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Nagasaki 25;  
 RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
 RA Yasuoka A., Oka S.;  
 RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
 RT infection: Prolonged HAV viremia and mild liver injury.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB038292; BAB11830.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 94  
 SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMGFSQAKISLF 20  
Db 74 VLPPPRKMGFSQAKISLF 93

## RESULT 11

Q9ENV1 PRELIMINARY; PRT; 94 AA.  
AC Q9ENV1;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.  
NCBI\_TaxID=12092;

[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Nagasaki 18;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
RT infection: Prolonged HAV viremia and mild liver injury.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB038286; BAB11824.1; -;  
FT NON\_TER 1 94  
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMGFSQAKISLF 20  
Db 74 VLPPPRKMGFSQAKISLF 93

## RESULT 12

Q9ENV9 PRELIMINARY; PRT; 94 AA.  
AC Q9ENV9;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.  
NCBI\_TaxID=12092;

[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Nagasaki 21;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
RT infection: Prolonged HAV viremia and mild liver injury.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB038288; BAB11826.1; -;  
FT NON\_TER 1 94  
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMGFSQAKISLF 20

Db 74 VLPPPRKMGFSQAKISLF 93

## RESULT 13

Q9ENV7 PRELIMINARY; PRT; 94 AA.  
AC Q9ENV7;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.  
NCBI\_TaxID=12092;

[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Nagasaki 06;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
RT infection: Prolonged HAV viremia and mild liver injury.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB038280; BAB11818.1; -;  
FT NON\_TER 1 94  
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMGFSQAKISLF 20  
Db 74 VLPPPRKMGFSQAKISLF 93

## RESULT 14

Q9ENV8 PRELIMINARY; PRT; 94 AA.  
AC Q9ENV8;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.  
NCBI\_TaxID=12092;

[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Nagasaki 04;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
RT infection: Prolonged HAV viremia and mild liver injury.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB038279; BAB11817.1; -;  
FT NON\_TER 1 94  
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 101; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPPRKMGFSQAKISLF 20  
Db 74 VLPPPRKMGFSQAKISLF 93

## RESULT 15



```

09ENV6
ID 09ENV6 PRELIMINARY: PRT: 94 AA.
AC 09ENV6;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID-12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 07;
RA Ida S.;
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038281; BAB11819.1; -.
NON_TER 1 1
SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

```

```

Query Match 100.0%; Score 101; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. NO. 1.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VLPPPRKMKGLFSQAKISLF 20
   |||
DB 74 VLPPPRKMKGLFSQAKISLF 93

```

Search completed: October 1, 2003, 10:02:36  
 Job time : 35.1765 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 ; Search time 12.7059 Seconds  
(without alignments)  
66.600 Million cell updates/sec

Title: US-09-171-432A-42

Perfect score: 101

Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PCrus.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/Backfilest.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	1091	6 5516630-2	Patent No. 5516630
2	101	100.0	2227	3 US-08-475-886-2	Sequence 2, Appl1
3	101	100.0	2227	3 US-08-475-886-4	Sequence 4, Appl1
4	101	100.0	2227	3 US-08-475-886-6	Sequence 6, Appl1
5	101	100.0	2227	3 US-08-397-232-2	Sequence 2, Appl1
6	101	100.0	2227	3 US-08-397-232-4	Sequence 4, Appl1
7	101	100.0	2227	3 US-09-171-387-2	Sequence 2, Appl1
8	101	100.0	2227	4 US-09-653-499-2	Sequence 2, Appl1
9	101	100.0	2227	4 US-09-653-499-4	Sequence 4, Appl1
10	101	100.0	2227	4 US-09-653-499-6	Sequence 6, Appl1
11	84	83.2	839	1 US-08-087-016-2	Sequence 2, Appl1
12	52.5	52.0	1025	2 US-08-304-309-2	Sequence 2, Appl1
13	52.5	52.0	1025	3 US-08-991-942-2	Sequence 2, Appl1
14	52.5	52.0	1025	3 US-09-138-103-2	Sequence 2, Appl1
15	52.5	52.0	1025	5 PCT-US95-04567-4	Sequence 4, Appl1
16	45.5	45.0	1025	5 PCT-US95-04567-2	Sequence 2, Appl1
17	45	44.6	23	1 US-07-800-364B-4	Sequence 4, Appl1
18	45	44.6	23	5 PCT-US91-03388-4	Sequence 4, Appl1
19	44	43.6	454	4 US-09-134-001C-4438	Sequence 4438, Ap
20	43.5	43.1	97	4 US-09-205-258-855	Sequence 855, App
21	42	41.6	345	4 US-09-252-991A-30726	Sequence 30726, A
22	42	41.6	1197	4 US-09-252-991A-30833	Sequence 30833, A
23	41.5	41.1	330	4 US-09-252-991A-23794	Sequence 23794, A
24	41	40.6	108	1 US-08-466-033-106	Sequence 106, App
25	41	40.6	108	2 US-08-444-733-106	Sequence 106, App
26	41	40.6	108	2 US-08-464-134-106	Sequence 106, App
27	41	40.6	108	2 US-08-461-361-106	Sequence 106, App

28	41	40.6	108	2 US-08-485-910-106	Sequence 106, App
29	41	40.6	108	5 PCT-US95-06266-89	Sequence 89, Appl
30	41	40.6	188	4 US-09-328-352-5612	Sequence 5612, Ap
31	41	40.6	616	1 US-08-638-911A-35	Sequence 35, Appl
32	41	40.6	853	1 US-08-638-911A-25	Sequence 25, Appl
33	41	40.6	853	1 US-08-638-911A-27	Sequence 27, Appl
34	41	40.6	853	1 US-08-638-911A-29	Sequence 29, Appl
35	41	40.6	853	1 US-08-638-911A-31	Sequence 31, Appl
36	41	40.6	981	4 US-09-252-991A-18616	Sequence 18616, A
37	41	40.6	2873	1 US-08-466-033-15	Sequence 15, Appl
38	41	40.6	2873	1 US-08-638-911A-2	Sequence 2, Appl1
39	41	40.6	2873	2 US-08-444-733-15	Sequence 15, Appl1
40	41	40.6	2873	2 US-08-464-134-15	Sequence 15, Appl1
41	41	40.6	2873	2 US-08-461-361-15	Sequence 15, Appl1
42	41	40.6	2873	2 US-08-485-910-15	Sequence 15, Appl1
43	41	40.6	2873	5 PCT-US95-06266-15	Sequence 15, Appl1
44	41	40.6	2905	4 US-08-466-260A-401	Sequence 401, App
45	41	40.6	2905	4 US-08-488-446-401	Sequence 401, App

## ALIGNMENTS

```
RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANIELLO, VINCENT R.;
; BAROUDY, BAHIGE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO: 2
; LENGTH: 1091
5516630-2

Query Match      100.0%: Score 101: DB 6: Length 1091:
Best Local Similarity 100.0%: Pred. No. 9.8e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VLPPPRKMKGLFSQAKISLF 20
DB      1060 VLPPPRKMKGLFSQAKISLF 1079

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN M
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
```

```
;; SEQ ID NO 2
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VLPPRRKMKGLFSQAKISLF 20
      |||
Db      823 VLPPRRKMKGLFSQAKISLF 842
```

```
RESULT 3
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
```

```
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US2
;; CURRENT APPLICATION NUMBER: US/08/475,886A
;; EARLIER FILING DATE: 1995-06-07
;; EARLIER APPLICATION NUMBER: 07/947,338
;; EARLIER FILING DATE: 1992-09-18
;; EARLIER APPLICATION NUMBER: 08/397,232
;; EARLIER FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VLPPRRKMKGLFSQAKISLF 20
      |||
Db      823 VLPPRRKMKGLFSQAKISLF 842
```

LT 4

```
;; Sequence 6, Application US/08475886A
;; Patent No. 6113912
;; GENERAL INFORMATION:
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US2
;; CURRENT APPLICATION NUMBER: US/08/475,886A
;; EARLIER FILING DATE: 1995-06-07
;; EARLIER APPLICATION NUMBER: 07/947,338
;; EARLIER FILING DATE: 1992-09-18
;; EARLIER APPLICATION NUMBER: 08/397,232
;; EARLIER FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 6
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VLPPRRKMKGLFSQAKISLF 20
      |||
Db      823 VLPPRRKMKGLFSQAKISLF 842
```

```
RESULT 5
US-08-397-232-2
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
```

```
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US1
;; CURRENT APPLICATION NUMBER: US/08/397,232A
;; EARLIER FILING DATE: 1995-04-17
;; EARLIER APPLICATION NUMBER: 07/947,338
;; EARLIER FILING DATE: 1992-09-18
;; EARLIER APPLICATION NUMBER: PCT/US93/08610
;; EARLIER FILING DATE: 1993-09-17
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 2
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 VLPPRRKMKGLFSQAKISLF 20
      |||
Db      823 VLPPRRKMKGLFSQAKISLF 842
```

```
RESULT 6
US-08-397-232-4
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
```

```
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US1
;; CURRENT APPLICATION NUMBER: US/08/397,232A
;; EARLIER FILING DATE: 1995-04-17
;; EARLIER APPLICATION NUMBER: 07/947,338
;; EARLIER FILING DATE: 1992-09-18
;; EARLIER APPLICATION NUMBER: PCT/US93/08610
;; EARLIER FILING DATE: 1993-09-17
;; NUMBER OF SEQ ID NOS: 4
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



PRIOR APPLICATION NUMBER: 08/475,886  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO: 6  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-09-653-499-6

Query Match 100.0%; Score 101; DB 4; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.1e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRKMKGLFSQAKISLF 20  
|||||  
823 VLPPRRKMKGLFSQAKISLF 842

RESULT 11  
US-08-087-016-2  
Sequence 2, Application US/08087016  
Patent No. 5430135  
GENERAL INFORMATION:  
APPLICANT: NAINAN, OMANA V.  
APPLICANT: MARGOLIS, HAROLD S.  
APPLICANT: ROBERTSON, BETTY H.  
APPLICANT: BRINTON, MARGO H.  
APPLICANT: EBERT, JAMES W.  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
STREET: 1615 L Street N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/087,016  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,828  
FILING DATE: 03-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: SCOTT, WATSON T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 839 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-087-016-2

Query Match 83.2%; Score 84; DB 1; Length 839;  
Best Local Similarity 83.3%; Pred. No. 5.5e-06;  
Matches 15; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRKMKGLFSQAKIS 18  
:|||||:|||||:|||||  
Db 822 ILPPRRKMKGLFSQAKIS 839

RESULT 12  
US-08-304-309-2  
Sequence 2, Application US/08304309  
Patent No. 5856454  
GENERAL INFORMATION:  
APPLICANT: GONZALEZ, Frank J.  
APPLICANT: FERNANDEZ-SALGUERO, Pedro  
TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/304,309  
FILING DATE: 09-SEP-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Smith, Timothy L.  
REGISTRATION NUMBER: 35,367  
REFERENCE/DOCKET NUMBER: 15280-210  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1025 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-304-309-2

Query Match 52.0%; Score 52.5; DB 2; Length 1025;  
Best Local Similarity 63.2%; Pred. No. 1.4;  
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 LPPRRKMKGLFSQAKISLF 20  
||||| : |||||  
Db 176 LPPRRKMKSEAVS-AKTALF 193

RESULT 13  
US-08-991-942-2  
Sequence 2, Application US/08991942  
Patent No. 6015673  
GENERAL INFORMATION:  
APPLICANT: GONZALEZ, Frank J.  
APPLICANT: FERNANDEZ-SALGUERO, Pedro  
TITLE OF INVENTION: CLONING AND EXPRESSION OF CDNA FOR HUMAN  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Stewart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/991,942
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/304,309
; FILING DATE: 09-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Timothy L.
; REGISTRATION NUMBER: 35,367
; REFERENCE/DOCKET NUMBER: 15280-210
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1025 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-991-942-2
;
; Query Match 52.0%; Score 52.5; DB 3; Length 1025;
; Best Local Similarity 63.2%; Pred. No. 1.4;
; Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
;
; QY 2 LPPPRKMGKLFQAKISLF 20
; DB 176 LPPPEKMESEAYS-AKIALF 193
;
; RESULT 14
; US-09-138-103-2
; Sequence 2, Application US/09138103A
; Patent No. 6232448
; GENERAL INFORMATION:
; APPLICANT: Yoshikubo, Takashi
; APPLICANT: Hasegawa, Masami
; TITLE OF INVENTION: Immunological Materials and Methods for Detecting
; FILE REFERENCE: 09/138,103 Yoshikubo, et al.
; CURRENT APPLICATION NUMBER: US/09/138,103A
; CURRENT FILING DATE: 1998-08-21
; EARLIER APPLICATION NUMBER: 97114630.3
; EARLIER FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-138-103-2
;
; Query Match 52.0%; Score 52.5; DB 3; Length 1025;
; Best Local Similarity 63.2%; Pred. No. 1.4;
; Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
;
; QY 2 LPPPRKMGKLFQAKISLF 20
; DB 176 LPPPEKMESEAYS-AKIALF 193
;
; RESULT 15
; PCT-US95-04567-4
; Sequence 4, Application PC/TUS9504567
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE

```

```

;
; TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: PatentIn Release #1.0, Version
; SOFTWARE: #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04567
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/227,357
; FILING DATE: 13-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: UOAB025P--
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1025 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-04567-4
;
; Query Match 52.0%; Score 52.5; DB 5; Length 1025;
; Best Local Similarity 63.2%; Pred. No. 1.4;
; Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
;
; QY 2 LPPPRKMGKLFQAKISLF 20
; DB 176 LPPPEKMESEAYS-AKIALF 193

```

Search completed: October 1, 2003, 10:06:29  
Job time : 13.7059 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 ; Search time 24.1176 Seconds  
(without alignments)  
131.201 Million cell updates/sec

Title: US-09-171-432a-42

Perfect score: 101  
Sequence: 1 VLPPPRKMKGLFSQAKISLF 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

al number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/FCI\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	352	15 US-10-272-459-45	Sequence 45, Appl
2	101	100.0	980	15 US-10-272-459-41	Sequence 41, Appl
3	101	100.0	2227	10 US-09-928-955-12	Sequence 12, Appl
4	101	100.0	2227	14 US-10-104-966-12	Sequence 12, Appl
5	101	100.0	2227	14 US-10-135-988-2	Sequence 2, Appl
6	101	100.0	2227	14 US-10-135-988-4	Sequence 4, Appl
7	101	100.0	2227	14 US-10-135-988-6	Sequence 6, Appl
8	70	69.3	835	15 US-10-272-459-40	Sequence 40, Appl
9	52.5	52.0	1025	9 US-09-834-886-2	Sequence 2, Appl
10	43.6	43.6	673	16 US-10-176-306-2	Sequence 2, Appl
11	43.5	43.1	97	12 US-09-933-767-855	Sequence 855, App
12	43.5	43.1	97	15 US-10-023-282-855	Sequence 855, App
13	42	41.6	79	15 US-10-106-698-4310	Sequence 4310, Ap
14	42	41.6	449	15 US-10-156-761-13588	Sequence 13588, A
15	42	41.6	2053	12 US-10-323-450-12	Sequence 12, Appl

16	42	41.6	2053	14 US-10-017-216-2	Sequence 2, Appl
17	41	40.6	2842	12 US-10-224-999A-3479	Sequence 3479, Ap
18	41	40.6	2905	8 US-08-424-550B-401	Sequence 401, App
19	41	40.6	2910	12 US-09-828-498-2	Sequence 2, Appl
20	40	39.6	166	10 US-09-851-026-11	Sequence 11, Appl
21	40	39.6	166	11 US-09-930-169-3	Sequence 3, Appl
22	40	39.6	272	11 US-09-934-455-140	Sequence 35, Appl
23	40	39.6	310	10 US-09-851-026-15	Sequence 36, Appl
24	40	39.6	312	10 US-09-851-026-36	Sequence 2, Appl
25	40	39.6	318	10 US-09-947-071-2	Sequence 472, App
26	40	39.6	320	15 US-10-102-806-472	Sequence 76, Appl
27	40	39.6	371	9 US-09-320-337-76	Sequence 4933, Ap
28	40	39.6	666	9 US-09-815-242-4933	Sequence 10663, A
29	40	39.6	857	11 US-09-991-936-1915	Sequence 1915, Ap
30	40	39.6	871	9 US-09-740-668A-32	Sequence 32, Appl
31	39.5	39.1	184	9 US-10-283-013-2	Sequence 2, Appl
32	39.5	39.1	390	10 US-09-903-340-195	Sequence 195, App
33	39.5	39.1	467	10 US-09-909-088B-195	Sequence 195, App
34	39.5	39.1	467	10 US-09-905-291A-195	Sequence 195, App
35	39.5	39.1	467	10 US-09-981-876-134	Sequence 195, App
36	39.5	39.1	467	10 US-09-902-853-195	Sequence 195, App
37	39.5	39.1	467	10 US-09-907-824-195	Sequence 195, App
38	39.5	39.1	467	10 US-09-907-841-195	Sequence 195, App
39	39.5	39.1	467	11 US-09-904-011-195	Sequence 195, App
40	39.5	39.1	467	11 US-09-906-742-195	Sequence 195, App
41	39.5	39.1	467	11 US-09-148-545-134	Sequence 134, App
42	39.5	39.1	467	11 US-09-906-838-195	Sequence 195, App
43	39.5	39.1	467	11 US-09-907-613-195	Sequence 195, App
44	39.5	39.1	467	11 US-09-907-613-195	Sequence 195, App
45	39.5	39.1	467	11 US-09-907-942-195	Sequence 195, App

## ALIGNMENTS

RESULT 1  
US-10-272-459-45  
; Sequence 45, Application US/10272459  
; Publication No. US20030124517A1  
; GENERAL INFORMATION:  
; APPLICANT: PICHANTRES, Sergio  
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT  
; FILE REFERENCE: PPI1955.002 / 2301-17955  
; CURRENT APPLICATION NUMBER: US/10/272,459  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein  
US-10-272-459-45  
  
Query Match 100.0%; Score 101; DB 15; Length 352;  
Best Local Similarity 100.0%; Pred. No. 1.3e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 VLPPPRKMKGLFSQAKISLF 20  
|||||  
Db 195 VLPPPRKMKGLFSQAKISLF 214  
  
RESULT 2  
US-10-272-459-41  
; Sequence 41, Application US/10272459  
; Publication No. US20030124517A1  
; GENERAL INFORMATION:  
; APPLICANT: PICHANTRES, Sergio  
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

```
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: P17955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41
```

```
Query Match          100.0%; Score 101; DB 15; Length 980;
Best Local Similarity 100.0%; Pred. No. 3.8e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 VLPPRRMKGLFSQAKISLF 20
      |||
Db      823 VLPPRRMKGLFSQAKISLF 842
```

```
RESULT 3
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12
```

```
Query Match          100.0%; Score 101; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 VLPPRRMKGLFSQAKISLF 20
      |||
Db      823 VLPPRRMKGLFSQAKISLF 842
```

```
RESULT 4
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
```

```
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12
```

```
Query Match          100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 VLPPRRMKGLFSQAKISLF 20
      |||
Db      823 VLPPRRMKGLFSQAKISLF 842
```

```
RESULT 5
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HW-175
US-10-135-988-2
```

```
Query Match          100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 VLPPRRMKGLFSQAKISLF 20
      |||
Db      823 VLPPRRMKGLFSQAKISLF 842
```

```
RESULT 6
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
```

;; PRIOR FILING DATE: 1995-03-10  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 4  
;; LENGTH: 2227  
;; TYPE: PRT  
;; ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
US-10-135-988-4

Query Match 100.0%; Score 101; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 9e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRKMKGLFSQAKISLF 20  
DB 823 VLPPRRKMKGLFSQAKISLF 842

RESULT 7  
US-10-135-988-6  
Sequence 6, Application US/10135988  
Publication No. US20020176869A1  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN M  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PORCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262US3  
CURRENT APPLICATION NUMBER: US/10/135,988  
PRIOR FILING DATE: 2002-04-29  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-10-135-988-6

Query Match 100.0%; Score 101; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 9e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VLPPRRKMKGLFSQAKISLF 20  
823 VLPPRRKMKGLFSQAKISLF 842

RESULT 8  
US-10-272-459-40  
Sequence 40, Application US/10272459  
Publication No. US20030124517A1  
GENERAL INFORMATION:  
APPLICANT: PICHANTES, Sergio  
TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT  
FILE REFERENCE: PPI7955.002 / 2301-17955  
CURRENT APPLICATION NUMBER: US/10/272,459  
PRIOR FILING DATE: 2002-10-15  
NUMBER OF SEQ ID NOS: 50  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 40  
LENGTH: 836  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: recombinant protein  
US-10-272-459-40

Query Match 69.3%; Score 70; DB 15; Length 836;  
Best Local Similarity 92.9%; Pred. No. 0.019;  
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VLPPRRKMKGLFSQ 14  
DB 823 VLPPRRKMKGLFSQ 836

RESULT 9  
US-09-854-886-2  
Sequence 2, Application US/09854886  
Patent No. US20020072080A1  
GENERAL INFORMATION:  
APPLICANT: Yoshikubo, Takashi  
APPLICANT: Hasegawa, Masami  
TITLE OF INVENTION: Immunological Material and Methods for Detecting  
FILE REFERENCE: 100554-32887  
CURRENT APPLICATION NUMBER: US/09/854,886  
PRIOR FILING DATE: 2001-05-14  
PRIOR APPLICATION NUMBER: 09/138,103  
PRIOR FILING DATE: 1998-08-21  
PRIOR APPLICATION NUMBER: 97114630.3  
PRIOR FILING DATE: 1997-08-22  
NUMBER OF SEQ ID NOS: 8  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 1025  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-854-886-2

Query Match 52.0%; Score 52.5; DB 9; Length 1025;  
Best Local Similarity 63.2%; Pred. No. 11;  
Matches 12; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 2 LPPRRKMKGLFSQAKISLF 20  
DB 176 LPPRRKMKGLFSQAKISLF 193

RESULT 10  
US-10-176-306-2  
Sequence 2, Application US/10176306  
Publication No. US20030130485A1  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel A.  
APPLICANT: Curtis, Rory A. J.  
APPLICANT: Glucksmann, Maria Alexandra  
APPLICANT: Bandaru, Rajasekhar  
APPLICANT: Kapeller-Libermann, Rosana  
TITLE OF INVENTION: NOVEL HUMAN GENES AND METHODS OF USE THEREOF  
FILE REFERENCE: 10448-195001  
CURRENT APPLICATION NUMBER: US/10/176,306  
PRIOR FILING DATE: 2002-06-20  
PRIOR APPLICATION NUMBER: 10/001,137  
PRIOR FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: PCT/US01/45291  
PRIOR FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/248,362  
PRIOR FILING DATE: 2000-11-14  
PRIOR APPLICATION NUMBER: 60/248,331  
PRIOR FILING DATE: 2000-11-14  
PRIOR APPLICATION NUMBER: 60/248,365  
PRIOR FILING DATE: 2000-11-14  
PRIOR APPLICATION NUMBER: 60/250,077  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: 60/250,327  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: 60/250,176  
PRIOR FILING DATE: 2000-11-30

;; PRIOR APPLICATION NUMBER: 10/023,617  
;; PRIOR FILING DATE: 2001-12-18  
;; PRIOR APPLICATION NUMBER: PCT/US01/49416  
;; PRIOR FILING DATE: 2001-12-18  
;; PRIOR APPLICATION NUMBER: 60/256,249  
;; PRIOR FILING DATE: 2000-12-18  
;; PRIOR APPLICATION NUMBER: 60/256,405  
;; PRIOR FILING DATE: 2000-12-18  
;; PRIOR APPLICATION NUMBER: 10/083,248  
;; PRIOR FILING DATE: 2001-10-22  
;; PRIOR APPLICATION NUMBER: PCT/US01/46717  
;; PRIOR FILING DATE: 2001-10-22  
;; PRIOR APPLICATION NUMBER: 60/242,324  
;; PRIOR FILING DATE: 2000-10-20  
;; PRIOR APPLICATION NUMBER: 60/242,518  
;; PRIOR FILING DATE: 2000-10-23  
;; PRIOR APPLICATION NUMBER: 60/241,989  
;; PRIOR FILING DATE: 2000-10-20  
;; NUMBER OF SEQ ID NOS: 86  
;; SOFTWARE: FASTSEQ for Windows Version 4.0  
;; SEQ ID NO 2  
;; LENGTH: 673  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-176-306-2

Query Match 43.6%; Score 44; DB 16; Length 673;  
Best local similarity 43.8%; Pred. No. 1.5e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

OY 3 PPPKMGKFSQAKIS 18  
||||:||||:|  
Db 507 PPPGKGSFSELT 522

RESULT 11  
US-09-933-767-855  
;; Sequence 855, Application US/09933767  
;; Publication No. US20030181692A1  
;; GENERAL INFORMATION:  
;; APPLICANT: NI et al.  
;; TITLE OF INVENTION: 207 Human Secreted Proteins  
;; FILE REFERENCE: P2007P2  
;; CURRENT APPLICATION NUMBER: US/09/933,767  
;; PRIOR FILING DATE: 2001-08-22  
;; PRIOR APPLICATION NUMBER: PCT/US01/05614  
;; PRIOR FILING DATE: 2001-02-21  
;; PRIOR APPLICATION NUMBER: 60/184,836  
;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: 60/193,170  
;; PRIOR FILING DATE: 2000-03-29  
;; PRIOR APPLICATION NUMBER: 09/205,258  
;; PRIOR FILING DATE: 1998-12-04  
;; PRIOR APPLICATION NUMBER: PCT/US98/11422  
;; PRIOR FILING DATE: 1998-06-04  
;; PRIOR APPLICATION NUMBER: 60/048,885  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/049,375  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,881  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,880  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,896  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/049,020  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,876  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,895  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,884  
;; PRIOR FILING DATE: 1997-06-06

;; PRIOR APPLICATION NUMBER: 60/048,894  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,971  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,964  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,882  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,899  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,893  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,900  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,901  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,892  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,915  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/049,019  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,970  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,972  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,916  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/049,373  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,875  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/049,374  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,917  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,949  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,974  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,883  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,897  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,898  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,962  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,963  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,877  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/048,878  
;; PRIOR FILING DATE: 1997-06-06  
;; PRIOR APPLICATION NUMBER: 60/058,054  
;; PRIOR FILING DATE: 1997-12-18  
;; PRIOR APPLICATION NUMBER: 60/068,064  
;; PRIOR FILING DATE: 1997-12-18  
;; PRIOR APPLICATION NUMBER: 60/068,053  
;; PRIOR FILING DATE: 1997-12-18  
;; PRIOR APPLICATION NUMBER: 60/070,923  
;; PRIOR FILING DATE: 1997-12-18  
;; PRIOR APPLICATION NUMBER: 60/073,160  
;; PRIOR FILING DATE: 1998-01-30  
;; PRIOR APPLICATION NUMBER: 60/073,159  
;; PRIOR FILING DATE: 1998-01-30  
;; PRIOR APPLICATION NUMBER: 60/073,165  
;; PRIOR FILING DATE: 1998-01-30  
;; PRIOR APPLICATION NUMBER: 60/073,164  
;; PRIOR FILING DATE: 1998-01-30  
;; PRIOR APPLICATION NUMBER: 60/085,925  
;; PRIOR FILING DATE: 1998-05-18  
;; PRIOR APPLICATION NUMBER: 60/085,921

```

; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,923
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/085,922
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/092,921
; PRIOR FILING DATE: 1998-07-15
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1245
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 855
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-933-767-855

Query Match          43.1%; Score 43.5; DB 12; Length 97;
Best Local Similarity 47.4%; Pred. No. 24;
Matches 9; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

2 LPPPRKMGVLA-IRLSLF 69
52 VPPPLKMGVLA-IRLSLF 69
Db

RESULT 12
US-10-023-282-855
; Sequence 855, Application US/10023282
; Publication No. US20030092893A1
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/10/023,282
; CURRENT FILING DATE: 2001-12-20
; EARLIER APPLICATION NUMBER: 08/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
```

```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 855
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-023-282-855

Query Match          43.1%; Score 43.5; DB 15; Length 97;
Best Local Similarity 47.4%; Pred. No. 24;
Matches 9; Conservative 5; Mismatches 4; Indels 1; Gaps 1;

2 LPPPRKMGVLA-IRLSLF 69
52 VPPPLKMGVLA-IRLSLF 69
Db

RESULT 13
US-10-106-698-4310
; Sequence 4310, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
```

```

; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4310
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (13)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (79)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-4310

```

```

Query Match 41.6%; Score 42; DB 15; Length 79;
Best Local Similarity 55.6%; Pred. No. 32;
Matches 10; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

```

```

QY 3 PPRKMGIFSQAISIF 20
DB 34 PPRKME-FRTASIRLF 49

```

```

RESULT 14
US-10-156-761-13588
; Sequence 13588, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13588
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13588

```

```

Query Match 41.6%; Score 42; DB 15; Length 449;
Best Local Similarity 44.4%; Pred. No. 2e+02;
Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1 VLPPRRKMGIFSQAIS 18
DB 381 LMAPRGVRLVFSKAGVT 398

```

```

RESULT 15
US-10-325-430-12
; Sequence 12, Application US/10325430
; Publication No. US20030153525A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc
; APPLICANT: Silos-Santiago, Immaculada
; APPLICANT: Rosenfeld, Julie Beth

```

```

; TITLE OF INVENTION: METHODS AND COMPOSITIONS IN TREATING
; TITLE OF INVENTION: PAIN AND PAINFUL DISORDERS USING 1465, 1587, 2146, 2207,
; FILE REFERENCE: MP101-294P1RNM
; CURRENT APPLICATION NUMBER: US/10/325,430
; CURRENT FILING DATE: 2002-12-19
; PRIOR FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2053
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-325-430-12

```

```

Query Match 41.6%; Score 42; DB 12; Length 2053;
Best Local Similarity 61.5%; Pred. No. 9.8e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 4 PPRKMGIFSQA 16
DB 1259 PAKKKGLFSRRK 1271

```

```

Search completed: October 1, 2003, 10:37:50
Job time : 24.1176 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 ; Search time 43.5294 Seconds  
(without alignments)  
72.928 Million cell updates/sec

Title: US-09-171-432a-43  
Perfect score: 101  
Sequence: 1 FSQAKISLFTYEHEIMKFS 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
al number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq.19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	20	18	AAW42926
2	101	100.0	21	22	AB69443
3	101	100.0	854	6	AA50287
4	101	100.0	1077	20	AAW95559
5	101	100.0	1091	14	AAK32426
6	101	100.0	2227	11	AAK05697
7	101	100.0	2227	18	AAW34074
8	101	100.0	2227	21	AA18607
9	101	100.0	2227	21	AA18608

10	101	100.0	2227	21	AA18609
11	101	100.0	2227	23	ABG31727
12	101	100.0	2227	23	ABG31728
13	101	100.0	2227	23	ABG31729
14	101	100.0	2227	23	AAE19899
15	101	100.0	2227	24	ABU08639
16	101	100.0	2227	24	ABU08640
17	101	100.0	2227	24	ABU08641
18	98	97.0	2227	7	AA60066
19	97	96.0	366	6	AA50230
20	97	96.0	993	6	AA50231
21	97	96.0	993	6	AA50231
22	49.5	49.0	20	19	AAW76539
23	47	46.5	518	20	AAV27340
24	47	46.5	816	20	AAV27339
25	47	46.5	846	23	ABP30554
26	47	46.5	847	23	ABP27030
27	47	46.5	847	23	ABP29773
28	43	42.6	171	23	ABP40360
29	42	41.6	20	18	AAW42925
30	42	41.6	20	18	AAW42925
31	42	41.6	125	23	ABW49486
32	42	41.6	172	23	ABP25691
33	42	41.6	503	22	ABG15762
34	41	40.6	144	22	ABW96238
35	41	40.6	173	22	AAW49279
36	41	40.6	238	22	ABW68889
37	41	40.6	2431	13	AAK25138
38	40	39.6	105	21	AAK19719
39	40	39.6	105	21	AAK61014
40	40	39.6	122	21	AAK41997
41	40	39.6	173	22	AAU29276
42	40	39.6	173	22	AAW49276
43	40	39.6	173	24	ABU71364
44	40	39.6	173	24	ABU65821
45	40	39.6	173	24	ABU65154

#### ALIGNMENTS

RESULT 1  
ID AAW42926 standard; peptide: 20 AA.  
XX AAW42926;  
AC AAW42926;  
XX 28-APR-1998 (first entry)  
DT 28-APR-1998 (first entry)  
XX Immunogenic Hepatitis A virus peptide YK-1319.  
DE Immunogenic Hepatitis A virus peptide YK-1319.  
XX Immunogenic peptide; immunogenic epitope; p2a protein;  
KW Immune response; antibody.  
KM Synthetic.  
XX Hepatitis A virus.  
OS WO9740147-A1.  
PN 30-OCT-1997.  
PD 18-APR-1997; 97WO-US06891.  
PF 19-APR-1996; 96US-0015644.  
PR (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA Fields HA, Khudaykov YE;  
PI WPI, 1997-535831/49.  
XX Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an  
XX Immune response to HAV in a mammal or to detect the presence of

PT antibodies against HAV in a mammal  
 PS Claim 18; Page 112; 140pp; English.  
 XX  
 CC Peptides AAW42922-30 are immunogenic peptides corresponding to  
 CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are  
 CC substantially similar to a portion of the amino acid sequence of the P2A  
 CC protein of HAV corresponding to amino acids 792-980. The present peptide  
 CC is derived from amino acids 834-853, and has a reactivity of 27.1% with  
 CC acute sera. Compositions containing the peptides can be used to induce an  
 CC immune response to HAV in a mammal. The peptides can also be used to  
 CC detect the presence of antibodies against HAV in mammalian serum. The  
 CC peptides can also be used to make an antibody against HAV by  
 CC administering the peptide to a mammal.  
 CC  
 XX Sequence 20 AA;  
 SQ  
 Query Match 100.0%; Score 101; DB 18; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-10;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1 FSOAKISLFYTEHEHMKFS 20  
 |||||  
 1 FSOAKISLFYTEHEHMKFS 20  
 RESULT 2  
 AAB69443  
 ID AAB69443 standard; Peptide; 21 AA.  
 XX  
 AC AAB69443;  
 XX  
 DT 20-APR-2001 (first entry)  
 XX  
 DE Synthetic HAV P2A peptide, SEQ ID NO: 43.  
 XX  
 KM Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;  
 KM antigen; major structural capsid polypeptide; HAV antibody detection.  
 XX  
 OS Hepatitis A virus.  
 OS Synthetic.  
 PN WO200105824-A2.  
 XX  
 PD 25-JAN-2001.  
 XX  
 PF 14-JUL-2000; 2000WO-US19267.  
 XX  
 PF 15-JUL-1999; 99US-0144412.  
 XX  
 (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Fields RA, Khudyakov YE;  
 DR WPI; 2001-112681/12.  
 XX  
 PT Synthetic peptides used as antigen sources for enzyme immunoassays  
 PT detecting anti-hepatitis A virus and as vaccines -  
 XX  
 PS Claim 13; Page 95; 130pp; English.  
 XX  
 CC The present sequence is one of a number of synthetic peptides which are  
 CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides  
 CC comprise antigenic epitopes of the major structural capsid polypeptides  
 CC or non-structural polypeptides of HAV with one or more glutamine  
 CC molecules at the carboxy end of the peptide. The peptides are used to  
 CC detect the presence of antibodies against HAV in mammalian serum, to  
 CC detect the presence of HAV in a human or animal through the binding of  
 CC the peptide to an antibody, to detect acute phase infection by detecting  
 CC IgM antibodies in mammalian serum and detecting convalescence in a  
 CC mammal. The peptides are used to detect or quantify HAV antibodies in  
 CC samples in clinical or research-based assays using immunoblotting,  
 CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,

CC tracking of radioactive or bioluminescent markers, chromatography or  
 CC electrophoresis. The peptides are used to induce an immune response to  
 CC HAV when administered to a human or animal. Glutamine at the carboxy  
 CC end of the peptides enhances the IgM antibody reactivity.  
 CC  
 XX Sequence 21 AA;  
 SQ  
 Query Match 100.0%; Score 101; DB 22; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-10;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1 FSOAKISLFYTEHEHMKFS 20  
 |||||  
 1 FSOAKISLFYTEHEHMKFS 20  
 RESULT 3  
 AAP50287  
 ID AAP50287 standard; Protein; 854 AA.  
 XX  
 AC AAP50287;  
 XX  
 DT 25-MAR-2003 (updated)  
 DT 30-NOV-1991 (first entry)  
 XX  
 DE Sequence encoded by hepatitis A virus (HAV) cDNA from near the  
 DE genome 5' terminus to the end of the area corresponding to the  
 DE capsid protein region of poliovirus RNA.  
 XX  
 KM Hepatitis A virus assay; antigen; antibody.  
 XX  
 OS Hepatitis A virus.  
 OS WO8501517-A.  
 PN 11-APR-1985.  
 PD 27-SEP-1984; 84WO-US01552.  
 PF 30-SEP-1983; 83US-0537911.  
 PR  
 XX (MASI ) MASSACHUSETTS INST TECHNOLOGY.  
 PA  
 PI Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;  
 PI Racanelli VR;  
 XX  
 DR WPI; 1985-098846/16.  
 DR N-PSDB: AAN50330.  
 XX  
 FT New hepatitis A virus cDNA - useful in assays for the virus and  
 FT for prodn. of the viral antigen and antibodies to it  
 XX  
 PS Example; Fig 7; 60pp; English.  
 XX  
 CC The inventors claim HAV cDNA and a method for producing it, whereby  
 CC large amts. can be obtd. economically. The cDNA is useful in the  
 CC assay for detection of HAV quickly and easily and with high  
 CC sensitivity and specificity. The HAV cDNA is also used in the prodn.  
 CC of HAV antigen or antibodies to it. The antibodies may be monoclonal.  
 CC (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 854 AA;  
 Query Match 100.0%; Score 101; DB 6; Length 854;  
 Best Local Similarity 100.0%; Pred. No. 5e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1 FSOAKISLFYTEHEHMKFS 20  
 |||||  
 834 FSOAKISLFYTEHEHMKFS 853  
 RESULT 4



```

AAW95559
ID AAW95559 standard; Protein; 1077 AA.
XX
AC AAW95559;
XX
XX
XX 28-APR-1999 (first entry)
DE A partial hepatitis A virus (HAV) protein.
XX
XX Hepatitis A virus protein; HAV; P2 region;
XX cell-culture-adapted HAV strain; infection; accelerated growth.
XX
XX Hepatitis A virus.
XX
XX US5849562-A.
XX
XX 15-DEC-1998.
XX
XX 06-JUN-1995; 95US-0468926.
XX
XX 06-NOV-1991; 91US-0788262.
XX 30-SEP-1983; 83US-0537911.
XX 27-SEP-1984; 84US-0654942.
XX 06-OCT-1988; 88US-0256135.
XX 06-JUN-1995; 95US-0468926.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Emerson SU, Purcell RH;
XX
XX WPI; 1999-094412/08.
XX N-PSDB; AAX01006.
XX
XX Chimeric hepatitis A virus strains - with P2 region from
XX cell-culture-adapted strain in wild-type genome
XX
XX Disclosure; Fig 7A-L; 36pp; English.
XX
XX The present sequence represents a partial hepatitis A virus (HAV)
XX protein. The specification describes a DNA construct consisting
XX of a wild-type HAV genome in which the P2 region is replaced by the
XX P2 region from a cell-culture-adapted HAV strain. The construct is
XX used to demonstrate that mutations in the P2 region of a
XX cell-culture-adapted HAV strain are sufficient for establishment of
XX infection and accelerated growth in cell culture.
XX
XX Sequence 1077 AA:
SQ
Query Match 100.0%; Score 101; DB 20; Length 1077;
Best Local Similarity 100.0%; Pred. No. 6.4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FSOAKISLFYTEHEIMKFS 20
DB 1057 FSOAKISLFYTEHEIMKFS 1076

```

```

FH Key Location/Qualifiers
FT Region 238..1091
FT /label= ORF
FT /note= "second putative initiation codon at
FT position 240"
FT Region 1..711
FT /note= "X's correspond to nonsense codons,
FT i.e. this region is not an ORF"
XX
XX USN7788262-N.
XX
XX 15-DEC-1992.
XX
XX 30-SEP-1983; 83US-0536911.
XX
XX 27-SEP-1984; 84US-0654942.
XX 06-OCT-1988; 88US-0256135.
XX 30-SEP-1983; 83US-0536911.
XX 06-NOV-1991; 91US-0788262.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
XX Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH, Racanelli VR;
XX Baroudy BM, Emerson SU;
XX
XX WPI; 1993-067429/08.
XX N-PSDB; AAQ36934.
XX
XX Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn.
XX of antigen and antibodies
XX
XX Disclosure; Fig 7; 65pp; English.
XX
XX HAV virion RNA was extracted from the livers of marmosets which had
XX been inoculated with HAV (the HAV had previously been passaged twice
XX in marmosets). The RNA was used to prepare ds cDNA clones by
XX standard methods. Clones contg. inserts which hybridised to RNA from
XX HAV-infected African Green Monkey kidney cells were selected for
XX further analysis. A 7.4kb restriction map (about 9% of the HAV
XX genome) was constructed from 5 overlapping inserts. The sequence of
XX the first 3.3kb (approx.) from the 5' terminus was determined. An
XX amino acid sequence was decoded from the entire clone and an open
XX reading frame was identified starting at position 238. A comparison
XX of the predicted HAV amino acid sequences with the known capsid
XX protein sequences of other picornaviruses (poliovirus, foot and
XX mouth disease virus and encephalomyelitis virus) revealed areas of
XX local homology.
XX
XX (Note: Revised entry submitted to correct the patent number format of
XX US Government-owned WIS applications to prevent clashes with ongoing US
XX granted patent numbers. For further information please visit the Derwent
XX web site at www.derwent.com/dwpi/updates/ntis\_us.html.)
XX
XX Sequence 1091 AA:
SQ
Query Match 100.0%; Score 101; DB 14; Length 1091;
Best Local Similarity 100.0%; Pred. No. 6.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FSOAKISLFYTEHEIMKFS 20
DB 1071 FSOAKISLFYTEHEIMKFS 1090

```

```

RESULT 5
AAW95559
ID AAW95559 standard; Protein; 1091 AA.
XX
XX AAW95559;
XX
XX 25-MAR-2003 (updated)
XX 17-DEC-2001 (updated)
XX 10-JUN-1993 (first entry)
XX
XX Translated from 5' region of Hepatitis A Virus genomic clone.
XX
XX HAV HM-175; chronic liver disease; picornavirus.
XX
XX Hepatitis A virus.
XX

```

```

RESULT 6
AAW95559
ID AAW95559 standard; Protein; 2227 AA.
XX
XX AAW95559;
XX
XX 25-MAR-2003 (updated)
XX 15-AUG-1990 (first entry)
XX

```

DE		Attenuated hepatitis A virus.
KW		Hepatitis A virus; vaccine; attenuated.
OS		Hepatitis A virus, strain HM-175.
FH	Key	Location/Qualifiers
FT	Region	1..23
FT	Region	/label-VP4 = 1A
FT	Region	24..245
FT	Region	/label-VP2 = 1B
FT	Region	246..491
FT	Region	/label-VP3 = 1C
FT	Region	492..791
FT	Region	/label-VP1 = 1D
FT	Region	792..980
FT	Region	/label=2A
FT	Region	981..1087
FT	Region	/label=3B
FT	Region	1088..1422
FT	Region	/label=2C
FT	Region	1423..1496
FT	Region	/label=3A
FT	Region	1497..1519
FT	Region	/label=3B - VPg
FT	Region	1520..1738
FT	Region	/label=3C
FT	Region	1739..2227
FT	Region	/label=3D
XX	US4894228-A.	
PD	16-JAN-1990.	
PF	12-JUL-1988;	88US-0217824.
PR	12-JUL-1988;	88US-0217824.
PR	12-JUL-1988;	88US-0652967.
XX	(USSH ) US DEPT HEALTH & HUMAN SERVICE.	
PA	Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstone SM,	
PI	Daeмер RJ, Gust ID;	
XX	WPI: 1990-075557/10.	
DR	N-PSDB; AA003512.	
XX	Vaccine against hepatitis A virus infection - comprises novel	
XX	attenuated hepatitis A virus strain.	
XX	Claim 1; Fig 1; 18pp; English.	
CC	The attenuated HAV is useful for inducing protective immunity against	
CC	HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by	
CC	several nucleotide changes distributed throughout the genome, is	
CC	attenuated for chimpanzees, elicits serum neutralising antibodies, and is	
CC	suitable for use as an HAV vaccine. It is noted that not all the changes	
CC	are necessary for attenuation and use as a vaccine.	
CC	(Updated on 25-MAR-2003 to correct PA field.)	
CC	(Updated on 25-MAR-2003 to correct PI field.)	
SQ	Sequence 2227 AA:	
Query Match	100.0%; Score 101; DB 11; Length 2227;	
Best Local Similarity	100.0%; Pred. No. 1.4e-07;	
Matches 20: Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	1 FSOAKISLFYTEHEIMKFS 20	
Db	834 FSOAKISLFYTEHEIMKFS 853	

AAW34074	standard; Protein; 2227 AA.
AAW34074	
27-APR-1998	(first entry)
Hepatitis A virus HM-175 protein sequence.	
HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus; infection; vaccine.	
Hepatitis A virus HM-175.	
Key	Location/Qualifiers
Protein	1..23
Protein	/label= VP4
Protein	24..245
Protein	/label= VP2
Protein	246..491
Protein	/label= VP3
Protein	492..791
Protein	/label= VP1
Protein	792..980
Protein	/label= 2A
Protein	981..1087
Protein	/label= 2B
Protein	1088..1432
Protein	/label= 2C
Protein	1423..1496
Protein	/label= 3A
Protein	1497..1519
Protein	/label= 3B
Protein	1520..1738
Protein	/label= 3C
Protein	1739..2227
Protein	/label= 3D
MO9740166-A2.	
30-OCT-1997.	
18-APR-1997;	97MO-US06506.
19-APR-1996;	96US-0015642.
(USSH ) US SEC DEPT HEALTH.	
(USSH ) US DEPT HEALTH & HUMAN SERVICES.	
Emerson SU, Purcell RH, Raychaudhuri G;	
WPI; 1997-535850/49.	
N-PSDB; AAT93023.	
Human attenuated HAV genome containing simian HAV 2C gene - useful as vaccines against HAV infection	
Disclosure; Fig 13A-D; 66pp; English.	
This protein sequence is encoded by the human hepatitis A virus (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain HAV/7 is obtained by passage of HM-175 in African Green Monkey kidney cells. A claimed DNA construct (1) comprises a genome of HAV, where the genome is a human attenuated HAV genome in which a region of the 2C gene has been replaced by a corresponding region from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The region of the 2C gene from AGM-27 contained in the construct preferably encodes amino acids 120-328 of the 2C protein, amino acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA transcript of (1); (2) a cell transfected with (1) or the RNA transcript of (1); (3) a HAV genome as above; (4) antibodies to the HAV of (3); and (5) a host cell containing the HAV of (3). (1) or its RNA transcript, can be used as a vaccine for preventing HAV in	

CC a mammal. (1) or the RNA transcript can also be used to stimulate  
 CC the production of protective antibodies in the mammal.

XX Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 18; Length 2227;

Best Local Similarity 100.0%; Pred. NO. 1.4e-07; Mismatches 0; Indels 0; Gaps 0;

Matches 20; Conservative 0;

1 FSOAKISLFYTEHEIMKFS 20

834 FSOAKISLFYTEHEIMKFS 853

RESULT 8

AAB18607

ID AAB18607 standard; Protein; 2227 AA.

XX AAB18607;

15-JAN-2001 (first entry)

DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.

XX Hepatitis A virus.

PN US6113912-A.

XX 05-SEP-2000.

PF 07-JUN-1995; 95US-0475886.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

DR WPI; 2000-586464/55.

DR N-PSDB; AAA75476.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell

PT infection, has modified genome compared to wild type

XX Disclosure; fig 6A-K; 72pp; English.

XX The present sequence is derived from a wild type hepatitis A virus

CC (HAV) strain HM-174. The sequence is modified to produce HAV which

CC are adapted to growth in the human fibroblast-like cell line MRC-5.

CC The HAV is able to propagate in MRC-5 cells and retain appropriate

CC attenuation. It is useful as a live vaccine for prophylaxis of

CC hepatitis A in humans and other primates.

XX Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. NO. 1.4e-07; Mismatches 0; Indels 0; Gaps 0;

1 FSOAKISLFYTEHEIMKFS 20

834 FSOAKISLFYTEHEIMKFS 853

AC AAB18608;

XX 15-JAN-2001 (first entry)

DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;

XX P-35 virus.

XX Hepatitis A virus.

PN US6113912-A.

XX 05-SEP-2000.

PF 07-JUN-1995; 95US-0475886.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

DR WPI; 2000-586464/55.

DR N-PSDB; AAA75477.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell

PT line useful as vaccine for protecting humans against hepatitis A virus

XX infection, has modified genome compared to wild type

XX Disclosure; Columns 67-78; 72pp; English.

XX The present sequence is derived from passage 35 of a wild type

CC hepatitis A virus (HAV) strain HM-174. The resulting virus is

CC designated P-35 virus. The sequence is modified to produce HAV which

CC are adapted to growth in the human fibroblast-like cell line MRC-5.

CC The HAV is able to propagate in MRC-5 cells and retain appropriate

CC attenuation. It is useful as a live vaccine for prophylaxis of

CC hepatitis A in humans and other primates.

XX Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 21; Length 2227;

Best Local Similarity 100.0%; Pred. NO. 1.4e-07; Mismatches 0; Indels 0; Gaps 0;

1 FSOAKISLFYTEHEIMKFS 20

834 FSOAKISLFYTEHEIMKFS 853

RESULT 10

AAB18609

ID AAB18609 standard; Protein; 2227 AA.

XX AAB18609;

15-JAN-2001 (first entry)

DE Amino acid sequence of live attenuated Hepatitis A virus 4380.

XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;

XX HAV 4380.

XX Hepatitis A virus.

PN US6113912-A.

XX 05-SEP-2000.

PF 07-JUN-1995; 95US-0475886.

```
XX 18-SEP-1992; 92US-0947338.
PR 17-SEP-1993; 93WO-US08610.
PR 10-MAR-1995; 95US-0397232.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;
DR N-PSDB; AAA75478.
XX
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell
PT line useful as vaccine for protecting humans against hepatitis A virus
PT infection, has modified genome compared to wild type
XX
PS Disclosure; Columns 93-104; 72pp; English.
XX
CC The present sequence is derived from a live attenuated hepatitis A
CC virus (HAV) of the invention, designated HAV 4380. The sequence is
CC produced by modifying wild type HAV strain HM-174. The HAV of the
CC invention are adapted to growth in the human fibroblast-like cell
CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain
CC appropriate attenuation. It is useful as a live vaccine for prophylaxis
CC of hepatitis A in humans and other primates.
SQ Sequence 2227 AA;
Query Match 100.0%; Score 101; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FSOAKISLFYTEHEIMKFS 20
DB 834 FSOAKISLFYTEHEIMKFS 853
RESULT 11
ABG31727
ID ABG31727 standard; Protein; 2227 AA.
XX
AC ABG31727;
XX
DT 29-NOV-2002 (first entry)
XX
DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.
XX
KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.
XX
KW Hepatitis A virus strain HM-175.
PN US6423318-B1.
XX
PD 23-JUL-2002.
XX
PF 31-AUG-2000; 2000US-0653499.
XX
PR 07-JUN-1995; 95US-0475886.
PR 17-SEP-1993; 93US-0397232.
PR 17-SEP-1993; 93WO-US08610.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
XX
DR MPI: 2002-680946/73.
DR N-PSDB; ABS52787.
XX
PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection -
XX
PS Disclosure; Fig 6; 71pp; English.
```

```
XX
CC The invention relates to a polynucleotide which encodes a hepatitis A
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC line). The polynucleotide is useful for preparing a vaccine against
CC hepatitis A virus infection. This sequence represents a hepatitis A virus
CC strain HM-175 polypeptide.
XX
SQ Sequence 2227 AA;
Query Match 100.0%; Score 101; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 FSOAKISLFYTEHEIMKFS 20
DB 834 FSOAKISLFYTEHEIMKFS 853
RESULT 12
ABG31728
ID ABG31728 standard; Protein; 2227 AA.
XX
AC ABG31728;
XX
DT 29-NOV-2002 (first entry)
XX
DE Hepatitis A virus mutant strain HM-175/7 (PHAV/7) polypeptide.
XX
KW Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;
KW virucide; mutant; PHAV/7; mutin.
XX
OS Hepatitis A virus strain HM-175.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT MISC-difference 963
FT MISC-difference 764 /label= wild-type Lys substituted by Arg
FT MISC-difference 764 /note= "Wild-type Glu substituted by Val"
FT MISC-difference 821
FT MISC-difference 1052 /note= "Wild-type Asn substituted by Ser"
FT MISC-difference 1052
FT MISC-difference 1062 /note= "Wild-type Ala substituted by Val"
FT MISC-difference 1062
FT MISC-difference 1118 /note= "Wild-type Gly substituted by Ala"
FT MISC-difference 1118
FT MISC-difference 1151 /note= "Wild-type Lys substituted by Met"
FT MISC-difference 1151
FT MISC-difference 1163 /note= "Wild-type Glu substituted by Lys"
FT MISC-difference 1163 /note= "Wild-type Phe substituted by Ser"
FT MISC-difference 1277 /note= "Wild-type Val substituted by Ile"
FT MISC-difference 1500 /note= "Wild-type His substituted by Tyr"
FT MISC-difference 1805 /note= "Wild-type Asp substituted by Asn"
FT MISC-difference 1930 /note= "Wild-type Ser substituted by Thr"
XX
PN US6423318-B1.
XX
PD 23-JUL-2002.
XX
PF 31-AUG-2000; 2000US-0653499.
XX
PR 07-JUN-1995; 95US-0475886.
PR 17-SEP-1993; 93US-0397232.
PR 17-SEP-1993; 93WO-US08610.
XX
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
```

PI Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;  
 XX WPI: 2002-680946/73.  
 DR N-PSDB; ABS52788.  
 XX  
 PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth  
 PT in MRC-5 cells, useful for preparing a vaccine against HAV infection  
 XX  
 PS Example 3; Column 67-78; 71pp; English.  
 CC The invention relates to a polynucleotide which encodes a hepatitis A  
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell  
 CC line). The polynucleotide is useful for preparing a vaccine against  
 CC hepatitis A virus infection. This sequence represents a hepatitis A virus  
 CC mutant strain HM-175/7 (pHAV/7) polypeptide.  
 XX  
 SO Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 23; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 FSOAKISLFYTEHEIMKFS 20  
 DB 834 FSOAKISLFYTEHEIMKFS 853

RESULT 13  
 ABG31729  
 ID ABG31729 standard; Protein; 2227 AA.  
 AC ABG31729;  
 XX  
 DT 29-NOV-2002 (first entry)  
 XX  
 DE Attenuated Hepatitis A virus (HAV) 4380 polypeptide.  
 XX  
 KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;  
 KW HAV 4380.  
 XX  
 OS Hepatitis A virus strain HM-175.  
 XX  
 PN US6423318-B1.  
 XX  
 PD 23-JUL-2002.  
 XX  
 PF 31-AUG-2000; 2000US-0653499.  
 XX  
 PR 07-JUN-1995; 95US-0475886.  
 PR 17-SEP-1993; 93US-0397232.  
 PR 17-SEP-1993; 93WO-US08610.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;  
 XX  
 DR WPI: 2002-680946/73.  
 DR N-PSDB; ABS52789.  
 XX  
 PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth  
 PT in MRC-5 cells, useful for preparing a vaccine against HAV infection  
 XX  
 PS Disclosure: Column 93-104; 71pp; English.  
 CC The invention relates to a polynucleotide which encodes a hepatitis A  
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell  
 CC line). The polynucleotide is useful for preparing a vaccine against  
 CC hepatitis A virus infection. This sequence represents an attenuated  
 CC hepatitis A virus 4830 polypeptide.  
 XX  
 SO Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 23; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20  
 DB 834 FSOAKISLFYTEHEIMKFS 853

## RESULT 14

AAE19899  
 ID AAE19899 standard; Protein; 2227 AA.

AC AAE19899;  
 XX

DT 18-JUN-2002 (first entry)  
 XX

DE Hepatitis A virus (HAV) protein.  
 XX

KW Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;  
 KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.

OS Hepatitis A virus.  
 XX

PN WO200213855-A2.  
 XX

PD 21-FEB-2002.  
 XX

PF 15-AUG-2001; 2001WO-IB01808.  
 XX

PR 17-AUG-2000; 2000US-225767P.  
 PR 29-AUG-2000; 2000US-229175P.  
 PR 03-NOV-2000; 2000US-0705547.

XX  
 PA (TRIP-) TRIPEP AB.  
 XX

PI Salberg M, Hultgren C;  
 XX

DR WPI: 2002-241837/29.  
 DR N-PSDB; AAD31766.  
 XX

XX  
 PT Vaccine compositions for treating and preventing disease, preferably  
 PT hepatitis C virus infection, comprises ribavirin and antigen that has  
 PT epitope present in hepatitis C virus  
 XX

PS Claim 11; Page 82-87; 120pp; English.  
 XX

CC The invention relates to a composition comprising ribavirin and an  
 CC antigen preferably non structural 3 protein (NS3)/4A fragment of  
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV  
 CC sequence. The composition is useful for enhancing an immune response to  
 CC a hepatitis C antigen in humans, domestic, sport or pet species and as  
 CC vaccines for treating and preventing HCV infections. The composition is  
 CC also useful for treating viral, bacterial, fungal diseases and cancer.  
 CC The present sequence is hepatitis A virus (HAV) protein.  
 XX

SO Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 23; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20  
 DB 834 FSOAKISLFYTEHEIMKFS 853

RESULT 15  
 ABU08639  
 ID ABU08639 standard; Protein; 2227 AA.  
 XX  
 AC ABU08639;  
 XX

DT	03-JUN-2003	(first entry)
XX		
DE	Wild type human hepatitis A virus strain HM-175.	
XX		
KW	Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;	
KV	vaccine; MRC-5 cell; hepatitis infection.	
XX		
OS	Hepatitis A virus strain HM-175.	
XX		
PN	US2002176869-A1.	
PD		
XX	28-NOV-2002.	
PE		
PR	29-APR-2002; 2002US-0135988.	
XX		
PR	07-JUN-1995; 95US-0475886.	
PR	31-AUG-2000; 2000US-0653499.	
PR	18-SEP-1992; 92US-0947338.	
PR	17-SEP-1993; 93MO-US08610.	
PR	17-APR-1995; 95US-0397232.	
XX		
PA	(FUNK/) FUNKHOUSER A W.	
PA	(EMER/) EMERSON S U.	
PA	(PURC/) PURCELL R H.	
PA	(DHON/) D'HOND E.	
XX		
PI	Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;	
DR	WPI: 2003-352605/33.	
DR	N-PDB: ABX93473.	
XX		
PT	New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,	
PT	useful in vaccines for protecting primates against hepatitis infection	
XX	and disease -	
PS		
PS	Disclosure: Fig 6; 70pp: English.	
XX		
CC	The invention describes a live hepatitis A virus (HAV) adapted to growth	
CC	in MRC-5 cells. The HAV and compositions comprising the HAV are useful as	
CC	a vaccine for protecting primates against hepatitis infection and	
CC	disease. This is the amino acid sequence of wild type human hepatitis A	
CC	virus strain HM-175.	
XX		
SQ	Sequence 2227 AA:	
	Query Match 100.0%; Score 101; DB 24; Length 2227;	
	Best Local Similarity 100.0%; Pred. NO. 1,4e-07;	
	Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
	1 FSOAKISLFYEHEHIMKES 20	
	834 FSOAKISLFYEHEHIMKES 853	

Search completed: October 1, 2003, 09:56:42  
Job time : 44.5294 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 12.8235 seconds

(without alignments)  
149.988 Million cell updates/sec

Title: US-09-171-432a-43

Perfect score: 101

Sequence: 1 FSOAKISLFTYEHEIMKFS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	1358	2 A03905	genome polyprotein
2	101	100.0	2227	1 GNNYHM	genome polyprotein
3	101	100.0	2227	1 GNNYHR	genome polyprotein
4	101	100.0	2227	1 GNNYMK	genome polyprotein
5	101	100.0	2227	1 GNNYHB	genome polyprotein
6	97	96.0	852	1 GNNYHA	genome polyprotein
7	88	87.1	2230	1 GNNYSA	genome polyprotein
8	47	46.5	378	2 F64300	formate dehydrogenase
9	43	42.6	325	2 H71869	hypothetical prote
10	43	42.6	410	2 T22413	hypothetical prote
11	42	41.6	125	2 AB1747	glycine cleavage s
12	42	41.6	125	2 A11377	glycine cleavage s
13	42	41.6	187	2 T18844	hypothetical prote
14	42	41.6	545	2 S10842	gene ND5 intron 2
15	42	41.6	1265	2 F88866	protein F5262.2 [1
16	41.5	41.1	288	2 S64243	hypothetical prote
17	41	40.6	144	2 E75128	glycine cleavage s
18	41	40.6	190	2 C90529	nuclease, lipoprot
19	41	40.6	209	2 S65300	hypothetical prote
20	41	40.6	304	2 C72324	homoserine O-succ
21	41	40.6	465	2 B97235	deacetylase/di
22	41	40.6	483	2 G84113	hypothetical prote
23	41	40.6	557	2 S58164	glucose-6-phosphat
24	41	40.6	820	2 T51510	hypothetical prote
25	41	40.6	1040	1 A38306	alpha-mannosidase
26	41	40.6	2431	1 MNWVSF	nonstructural poly
27	40.5	40.1	251	2 T13104	minor tail protein
28	40.5	40.1	419	2 T25577	hypothetical prote
29	40	39.6	142	2 C70059	hypothetical prote

30	40	39.6	203	2 S63633	hypothetical prote
31	40	39.6	256	2 T45895	hypothetical prote
32	40	39.6	298	2 A89009	protein T27CA.1 [1
33	40	39.6	333	2 T13866	hypothetical prote
34	40	39.6	450	2 C82204	heat shock protein
35	40	39.6	523	2 D82631	probable sulfate t
36	39.5	39.1	109	2 S74054	hypothetical prote
37	39.5	39.1	260	2 B71353	probable D,D-carbo
38	39.5	39.1	410	2 T23264	hypothetical prote
39	39.5	39.1	507	2 B89792	hypothetical prote
40	39	38.6	129	2 F72075	hypothetical prote
41	39	38.6	129	2 G86547	hypothetical prote
42	39	38.6	145	2 F83247	hypothetical prote
43	39	38.6	152	1 C69202	conserved hypothet
44	39	38.6	167	2 F83845	hypothetical prote
45	39	38.6	171	2 T18167	hypothetical prote

## ALIGNMENTS

### RESULT 1

A03905

genome polyprotein (version 2) - human hepatitis A virus (fragments)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; prote

C:Species: human hepatitis A virus

C:Date: 31-Mar-1991 #sequence\_revision 31-Mar-1991 #text\_change 15-Nov-1996

R:Baroudy, B.M.; Ticehurst, J.R.; Miele, T.A.; Matzel Jr., J.V.; Purcell, R.H.; Feins

Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985

A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and R

A:Reference number: A03905; MUID:85166289; PMID:2984684

A:Accession: A03905

A:Molecule type: genomic RNA

A:Residues: 1-1358 <BAR>

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; polyprotein

F:1-245/Product: coat protein 1A #status predicted <C1A>

F:246-491/Product: coat protein 1B #status predicted <C1B>

F:492-836/Product: coat protein 1C #status predicted <C1C>

F:837-854/Product: core protein 2A (fragment) #status predicted <C2A>

F:855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

### Query Match

Best Local Similarity 100.0%; Score 101; DB 2; Length 1358;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFTYEHEIMKFS 20

DB 834 FSOAKISLFTYEHEIMKFS 853

### RESULT 2

GNNYHM genome polyprotein - human hepatitis A virus (strain HM-175, wild type)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core

B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 16-Jul-1999

C:Accession: A25981

R:Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with

A:Reference number: A25981; MUID:87061253; PMID:3023706

A:Accession: A25981

A:Molecule type: genomic RNA

A:Residues: 1-2227 <GENO>

A:Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA5465.1; PID:g329583

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-23/Product: coat protein 1A #status predicted <VP4>

F:24-245/Product: coat protein 1B #status predicted <VP2>





C:Accession: A03904  
R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra, S.W.  
J. Virol. 54, 247-255, 1985  
A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.  
A:Reference number: A03904; MUID:85185648; PMID:2985793  
A:Accession: A03904  
A:Molecule type: genomic RNA  
A:Residues: 1-852 <LIN>  
A:Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593  
C:Superfamily: hepatitis A virus genome polypeptide  
C:Keywords: coat protein; core protein; polypeptide  
F:1-245/Product: coat protein 1A #status predicted <C1A>  
F:246-491/Product: coat protein 1B #status predicted <C1B>  
F:492-836/Product: coat protein 1C #status predicted <C1C>  
F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 96.0%; Score 97; DB 1; Length 852;  
Best Local Similarity 100.0%; Pred. No. 4.7e-08;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 FSOAKISLFYTEHEIMKRF 19  
DB 834 FSOAKISLFYTEHEIMKRF 852  
|||||

RESULT 7  
GNNYSA  
genome polypeptide - simian hepatitis A virus (strain AGM-27)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro  
C:Species: simian hepatitis A virus  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jun-2000  
C:Accession: A30470; S04885; S03965  
R:Tsarev, S.A.  
submitted to JIPID, April 1991  
A:Reference number: A30470  
A:Accession: A30470  
A:Molecule type: genomic RNA  
A:Residues: 1-2230 <TSA>  
A:Cross-references: GB:D00924; NID:g222597; PIDN:BA00766.1; PID:g222598  
R:Tsarev, S.A.; Emerson, S.U.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.  
J. Gen. Virol. 72, 1677-1683, 1991  
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure an  
A:Reference number: J01080; MUID:91111420; PMID:1649901  
A:Contents: annotation  
A:Note: neither amino acid nor nucleotide sequence is given  
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik  
submitted to the EMBL Data Library, May 1989  
A:Reference number: S04885  
A:Accession: S04885  
A:Molecule type: genomic RNA  
A:Residues: 1750-2164 <BAL1>  
A:Cross-references: EMBL:X15461; NID:g61971; PIDN:CA33490.1; PID:g930268  
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizhik  
FEBS Lett. 247, 425-428, 1989  
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian he  
A:Reference number: S03965; MUID:8932168; PMID:2541023  
A:Accession: S03965  
A:Molecule type: genomic RNA  
A:Residues: 1960-2164 <BAL2>  
A:Cross-references: EMBL:X15461  
C:Superfamily: hepatitis A virus genome polypeptide  
C:Keywords: coat protein; core protein; polypeptide  
F:1-27/Product: coat protein 1A #status predicted <C1A>  
F:28-249/Product: coat protein 1B #status predicted <C1B>  
F:250-495/Product: coat protein 1C #status predicted <C1C>  
F:496-795/Product: coat protein 1D #status predicted <C1D>  
F:796-984/Product: core protein 2A #status predicted <C2A>  
F:985-1091/Product: core protein 2B #status predicted <C2B>  
F:1092-1426/Product: core protein 2C #status predicted <C2C>  
F:1427-1498/Product: core protein 3A #status predicted <C3A>  
F:1499-1521/Product: core protein 3B #status predicted <C3B>  
F:1522-1741/Product: protein 3C #status predicted <C3C>  
F:1742-2230/Product: protein 3D #status predicted <C3D>

Query Match 87.1%; Score 88; DB 1; Length 2230;  
Best Local Similarity 85.0%; Pred. No. 4.6e-06;  
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKRF 20  
DB 838 FSOAKISLFYTEHEIMKRF 857  
|||||

RESULT 8  
F64300  
formate dehydrogenase (EC 1.2.1.2) alpha chain - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: F64300  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak  
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weissbrock, K.G.; Merrick, J.M.; Glodek,  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Moese  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc  
A:Reference number: A64300; MUID:96337999; PMID:8688087  
A:Accession: F64300  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-378 <BUL>  
A:Cross-references: GB:U67459; GB:L77117; NID:g2826236; PIDN:AAB97987.1; PID:g1498763  
C:Genetics:  
A:Map position: REV8474-7338  
C:Keywords: oxidoreductase

Query Match 46.5%; Score 47; DB 2; Length 378;  
Best Local Similarity 47.1%; Pred. No. 4.7;  
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 ARISLFYTEHEIMKRF 20  
DB 174 SKWTIPTEKEIKLKN 190  
|||||

RESULT 9  
H71869  
hypothetical protein jhp0940 - Helicobacter pylori (strain J99)  
C:Species: Helicobacter pylori  
A:Variety: strain J99  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 04-Mar-2000  
C:Accession: H71869  
R:Alm, R.A.; Ling, L.S.-L.; Moir, D.T.; King, B.L.; Brown, E.D.; Dolg, P.C.; Smith, D.  
; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F  
Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p  
A:Reference number: A71800; MUID:99120557; PMID:9923682  
A:Accession: H71869  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-325 <ARN>  
A:Cross-references: GB:AE001523; GB:AE001439; NID:g4155513; PIDN:AAD06516.1; PID:g415  
A:Experimental source: strain J99  
C:Superfamily: Helicobacter pylori hypothetical protein jhp0940  
C:Genetics:  
A:Gene: jhp0940

Query Match 42.6%; Score 43; DB 2; Length 325;  
Best Local Similarity 60.0%; Pred. No. 19;  
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 KISLFYTEHEIMKRF 19  
DB 24 KISLFYNNELNWKVF 38  
|||||

RESULT 10





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 ; Search time 6.58824 Seconds

(without alignments)  
142.760 Million cell updates/sec

Title: US-09-171-432a-43

Perfect score: 101  
Sequence: 1 FSOAKISLFTYEHEIMKFS 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

al number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	2227	1	P08617 hepatitis a
2	101	100.0	2227	1	P06441 hepatitis a
3	101	100.0	2227	1	P13901 hepatitis a
4	97	96.0	852	1	P06442 hepatitis a
5	93	92.1	2226	1	P26580 hepatitis a
6	93	92.1	2226	1	P06442 hepatitis a
7	88	87.1	2230	1	P26582 hepatitis a
8	88	87.1	2230	1	P14553 simian hepa
9	47	46.5	378	1	P08314 methanococ
10	44	43.6	509	1	O20585 caenorhabd
11	43	42.6	410	1	PSD6_CAEEL
12	43	42.6	506	1	MATK_LATRI
13	43	42.6	722	1	YC06_KLEPN
14	42	41.6	125	1	GCSH_LISTIN
15	42	41.6	125	1	GCSH_LISTMO
16	41.5	41.1	288	1	NIF3_YEAST
17	41	40.6	138	1	GCSH_PYRAB
18	41	40.6	557	1	META_THEMA
19	41	40.6	557	1	G6P1_ACICA
20	41	40.6	1039	1	M2C1_MOUSE
21	41	40.6	1040	1	M2C1_RAT
22	41	40.6	2431	1	POLN_SFV
23	40	39.6	134	1	GCSH_PYRPU
24	40	39.6	418	1	B2AR_BOVIN
25	40	39.6	5147	1	PCLO_HUMAN
26	39.5	39.1	347	1	HTRPX_PYRAE
27	39	38.6	306	1	MK16_YEAST
28	39	38.6	380	1	TRPD_YEAST
29	39	38.6	534	1	UD15_HUMAN
30	39	38.6	634	1	YCX3_EUGAR
31	39	38.6	707	1	HLVB_ECOLI
32	39	38.6	707	1	HLVB_ECOLI
33	39	38.6	707	1	HLVB_PROVO

34	39	38.6	837	1	LDLR_RABIT	P20063 oryctolagus
35	39	38.6	860	1	LDLR_HUMAN	P01130 homo sapien
36	39	38.6	892	1	LDL2_XENLA	O99088 xenopus lae
37	39	38.6	909	1	LDL1_XENLA	O99087 xenopus lae
38	38	37.6	48	1	Y234_HAETN	P43968 haemophilus
39	38	37.6	138	1	GCSH_PYRHO	O59049 pyrococcus
40	38	37.6	265	1	CXBA_BACTW	O45790 bacillus th
41	38	37.6	317	1	RFAC_SALTU	P26469 salmonella
42	38	37.6	319	1	RFAC_ECOLI	P24173 escherichia
43	38	37.6	377	1	MFC3_CHVNI	P10835 chlorella v
44	38	37.6	396	1	REPA_BACST	P13862 bacillus su
45	38	37.6	400	1	ACH1_BOMMO	O03383 bombyx mori

## ALIGNMENTS

RESULT 1  
ID POLG\_HPVAH STANDARD; PRT; 2227 AA.  
AC P08617; P06443; O81082;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Genome polypeptide [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
DE P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain HM-175).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.  
OX NCBI\_Taxid=12098;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Wild type;  
RX MEDLINE=87061253; PubMed=3023706;  
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A., Baroudy B.M.;  
RT "Complete nucleotide sequence of wild-type hepatitis A virus: comparison with different strains of hepatitis A virus and other picornaviruses.";  
RT J. Virol. 61:50-59(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Attenuated;  
RX MEDLINE=87175701; PubMed=3031686;  
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Feinstein S.M., Purcell R.H.;  
RT "Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with wild-type virus.";  
RT Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).  
RN [3]  
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.  
RX MEDLINE=85166289; PubMed=2984684;  
RA Baroudy B.M., Ticehurst J.R., Miele T.A., Matzel J.V., Jr., Purcell R.H., Feinstein S.M.;  
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA polymerase.";  
RT Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + [RNA](n).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- MISCELLANEOUS: STRAIN HM-175/7 MR-5 IS ATTENUATED STRAIN DERIVED FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT SHOWN.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M1414; AAA45475.1; -  
DR EMBL: M14707; AAA45465.1; -  
DR EMBL: M14707; AAA45465.1; ALT-INIT.  
DR EMBL: M16632; AAA45471.1; -  
DR PIR: A25981; GNNYHM.  
DR PIR: A94149; GNNYMK.  
DR PDB: 1HAV; 23-DEC-96.  
DR MEROPS; C03.005; -  
DR InterPro: IPR004004; Calic1.pol.hel.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA.pol\_DS\_PS.  
DR InterPro: IPR001205; RNA.pol\_P3D.  
DR InterPro: IPR007094; RNA.pol\_PSV1r.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
DR PRINTS; PR00918; CALICIVIRUSNS.  
DR Polyprotein; Coat protein; Core protein; Transferase;  
RNA-directed RNA polymerase; Hydroxylase; Thiol protease; 3D-structure.

FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 836  
FT CHAIN 837 980  
FT CHAIN 981 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1496  
FT CHAIN 1497 1519  
FT CHAIN 1520 1738  
FT CHAIN 1739 2227  
FT CHAIN 77 77  
FT VARIANT 764 764  
FT VARIANT 821 821  
FT VARIANT 1052 1052  
FT VARIANT 1062 1062  
FT VARIANT 1118 1118  
FT VARIANT 1151 1151  
FT VARIANT 1163 1163  
FT VARIANT 1277 1277  
FT VARIANT 1500 1500  
FT VARIANT 1805 1805  
FT VARIANT 1930 1930  
SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AEB740A6 CRC64;

Query Match Best Local Similarity 100.0%; Score 101; DB 1; Length 2227;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFTYEEHEIMKFS 20  
Db 834 FSOAKISLFTYEEHEIMKFS 853

RESULT 2  
ID POLG\_HPAVL STANDARD: PRT: 2227 AA.  
AC P06441;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins  
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
P3D (EC 2.7.7.48)].  
DE P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain LA).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatitis A virus (strain LA).  
OC Hepatitis A virus (strain LA).  
OC NCBI\_Taxid=12099;  
OX NCBI\_Taxid=12099;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=65190549; PubMed=2986127;  
RA Natarajan R., Caput D., Gee W.W., Potter S.J., Renard A.,  
RA Merryweather J., van Nest G., Dina D.;  
RT "Primary structure and gene organization of human hepatitis A virus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +  
CC [RNA](N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: K02990; AAA45472.1; -  
DR PIR: A03903; GNNYHR.  
DR MEROPS; C03.005; -  
DR InterPro: IPR004004; Calic1.pol.hel.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA.pol\_DS\_PS.  
DR InterPro: IPR001205; RNA.pol\_P3D.  
DR InterPro: IPR007094; RNA.pol\_PSV1r.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
DR PRINTS; PR00918; CALICIVIRUSNS.  
DR Polyprotein; Coat protein; Core protein; Transferase;  
RNA-directed RNA polymerase; Hydroxylase; Thiol protease.

FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 836  
FT CHAIN 837 980  
FT CHAIN 981 1076  
FT CHAIN 1077 1422  
FT CHAIN 1423 1484  
FT CHAIN 1485 1507  
FT CHAIN 1508 1678  
FT CHAIN 1679 2227  
SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match Best Local Similarity 100.0%; Score 101; DB 1; Length 2227;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFTYEEHEIMKFS 20  
Db 834 FSOAKISLFTYEEHEIMKFS 853

RESULT 3  
ID POLG\_HPAVL STANDARD: PRT: 2227 AA.  
AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;  
AC Q81090; Q81091; Q81092; Q81093;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins  
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
P3D (EC 2.7.7.48)].  
DE P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain MBB).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatitis A virus (strain MBB).  
OC Hepatitis A virus (strain MBB).  
OC NCBI\_Taxid=12100;  
OX NCBI\_Taxid=12100;  
RN [1]  
RP SEQUENCE FROM N.A.

```

RX MEDLINE-86045071; PubMed-2823500;
RA Paul A.V., Tada H., der Helm R., Wissel T., Klein R., Wimmer E.,
RT Deinhardt F.;
RT "The entire nucleotide sequence of the genome of human hepatitis A
virus (isolate MBB)".
RL Virus Res. 8:153-171(1987).
CC
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
(RNA)(N).
CC
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
-1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERANEAN REGION.
CC
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC
CC
CC EMBL: M20273; AAA45474.1; -.
CC
CC DR MEROPS: C03.005; -.
CC
CC DR InterPro: IPR000605; RNA_helicase.
CC
CC DR InterPro: IPR007095; RNA_pol_DS_PS.
CC
CC DR InterPro: IPR007094; RNA_pol_PSVir.
CC
CC DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC
CC DR Pfam: PF00910; RNA_helicase; 1.
CC
CC KW Polypeptide; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
CC
CC FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
CC
CC FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
CC
CC FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
CC
CC FT CHAIN 492 836 COAT PROTEIN VP1 (PID).
CC
CC FT CHAIN 837 980 CORE PROTEIN P2A.
CC
CC FT CHAIN 981 1087 CORE PROTEIN P2B.
CC
CC FT CHAIN 1088 1422 CORE PROTEIN P2C.
CC
CC FT CHAIN 1423 1496 PROBABLE PROTEIN P3A.
CC
CC FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.
CC
CC FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.
CC
CC FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.
CC
CC FT CHAIN 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;
SQ
SQ SEQUENCE

```

```

Query Match 100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.2e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 834 FSOAKISLFYTEHEIMKFS 853

```

```

RESULT 4
POLG_HPAVC STANDARD: PRT; 852 AA.
ID POLG_HPAVC 083741; 083742;
AC 01-JAN-1988 (Rel. 06; Created)
DT 01-JAN-1988 (Rel. 06; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core protein
P2A] (Fragment).
OS Hepatitis A virus (strain CR326).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12097;
RN SEQUENCE FROM N.A.
RP MEDLINE=85185648; PubMed=2985793;
RA Linemeyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V.,

```

```

RA Young A., Mitra S.W.;
RT "Molecular cloning and partial sequencing of hepatitis A viral cDNA.";
RL J. Virol. 54:247-255(1985).
CC
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC
CC
CC EMBL: M10033; AAA45470.1; -.
CC
CC DR PIR: A03904; GNNYHA.
CC
CC KW Polypeptide; Coat protein; Core protein.
CC
CC FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
CC
CC FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
CC
CC FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
CC
CC FT CHAIN 492 836 COAT PROTEIN VP1 (PID).
CC
CC FT CHAIN 837 >852 CORE PROTEIN P2A.
CC
CC FT NON_TER 852 852
SQ SEQUENCE 852 AA; 95563 MW; 73D3ED0AD532820E CRC64;

```

```

Query Match 96.0%; Score 97; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 FSOAKISLFYTEHEIMKFS 19
DB 834 FSOAKISLFYTEHEIMKFS 852

```

```

RESULT 5
POLG_HPAV2 STANDARD: PRT; 2226 AA.
ID POLG_HPAV2
AC P26580;
DT 01-AUG-1992 (Rel. 23; Created)
DT 01-AUG-1992 (Rel. 23; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12094;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Cromeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
recombination."
RT variants arising during persistent infection: evidence for genetic
recombination."
RL J. Virol. 65:2056-2065(1991).
CC
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
(RNA)(N).
CC
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: M59810; AAA45468.1; -  
DR MEROPS: C03.005; -  
DR InterPro: IPR004004; Calic1\_pol\_hel.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR007094; RNA\_pol\_PSVLr.  
DR Pfam: PF00680; RNA\_dep\_RNA\_pol.1.  
DR Pfam: PF00910; RNA\_helicase.1.  
DR PRINTS: PR00918; CALICIVIRUSNS.  
KW Polypeptide; Coat protein; Core protein; Transferase;  
RNA-directed RNA polymerase; Hydrolase; Ribonuclease;  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 794  
FT CHAIN 795 900  
FT CHAIN 901 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1495  
FT CHAIN 1496 1518  
FT CHAIN 1519 1737  
FT CHAIN 1738 2226  
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;

Query Match 92.1%; Score 93; DB 1; Length 2226;  
Best Local Similarity 90.0%; Pred. No. 2.6e-07;  
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEHMKFS 20  
DB 834 FSOANISLFYTEHEHMKFS 853  
||||| ||||||| |||||

RESULT 6  
POLG\_HPAV4 STANDARD; PRT; 2226 AA.  
ID POLG\_HPAV4  
AC P26581;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins  
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain 43c).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatitis A virus;  
OC NCBI\_TaxID=12095;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91162758; PubMed=1705995;  
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,  
RA Cromeans T., Jansen R.W.,  
RT Antigenic and genetic variation in cytopathic hepatitis A virus  
RT variants arising during persistent infection: evidence for genetic  
RT recombination.  
RL J. Virol. 65:2056-2065(1991).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC (RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PFM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: M59809; AAA45469.1; -  
DR MEROPS: C03.005; -  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR007094; RNA\_pol\_PSVLr.  
DR Pfam: PF00680; RNA\_dep\_RNA\_pol.1.  
DR Pfam: PF00910; RNA\_helicase.1.  
KW Polypeptide; Coat protein; Core protein; Transferase;  
RNA-directed RNA polymerase; Hydrolase; Ribonuclease;  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 794  
FT CHAIN 795 900  
FT CHAIN 901 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1495  
FT CHAIN 1496 1518  
FT CHAIN 1519 1737  
FT CHAIN 1738 2226  
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B09BF75 CRC64;

Query Match 92.1%; Score 93; DB 1; Length 2226;  
Best Local Similarity 90.0%; Pred. No. 2.6e-07;  
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEHMKFS 20  
DB 834 FSOANISLFYTEHEHMKFS 853  
||||| ||||||| |||||

RESULT 7  
POLG\_HPAV8 STANDARD; PRT; 2226 AA.  
ID POLG\_HPAV8  
AC P26582;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins  
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain 18f).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatitis A virus;  
OC NCBI\_TaxID=12096;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91162758; PubMed=1705995;  
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,  
RA Cromeans T., Jansen R.W.,  
RT Antigenic and genetic variation in cytopathic hepatitis A virus  
RT variants arising during persistent infection: evidence for genetic  
RT recombination.  
RL J. Virol. 65:2056-2065(1991).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC (RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PFM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



```
CC -----
DR EMBL: M59808; AAA45467.1;
DR PDB: 1OAT; 15-MAY-00.
DR MEROPS: C03.005;
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVL.
DR Pfam: PF00680; RNA_dep_RNA_pol.1.
DR Pfam: PF00910; RNA_helicase.1.
DR Polyprotein; Coat protein; Core protein; Transferase;
KM RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.
ET CHAIN 1 23
ET CHAIN 24 245
ET CHAIN 246 491
ET CHAIN 492 794
ET CHAIN 795 900
ET CHAIN 901 1087
ET CHAIN 1088 1422
ET CHAIN 1423 1495
ET CHAIN 1496 1518
ET CHAIN 1519 1737
ET CHAIN 1738 2226
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8DB8 CRC64;

Query Match 92.1%; Score 93; DB 1; Length 2226;
Best Local Similarity 90.0%; Pred. No. 2.6e-07;
Matches 18; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEHMKFS 20
DB 834 FSOAKISLFYTEHEHMKFS 853

RESULT 8
POLG_HPAVS
ID FOLG_HPAVS STANDARD: PRT; 2230 AA.
AC P1453;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_Taxid=12102;
XX [1]
XX SEQUENCE FROM N.A.
RX MEDLINE-91311420; PubMed-1649901;
RA Tsarev S.A., Emerson S.O., Balayan M.S., Ticehurst J.R.,
RA Purcell R.H.;
RA "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome
RT structure and growth in cell culture with other HAV strains."
RT J. Gen. Virol. 72:1677-1683(1991).
RN [12]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE-89232168; PubMed-2541023;
RA Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chiznikov V.E., Blinov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses."
RL FEBS Lett. 247:425-428(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: D00924; BAA00766.1;
DR EMBL: X15461; CAA33490.1;
DR PIR: A30470; GNNYSA.
DR MEROPS: C03.005;
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVL.
DR Pfam: PF00680; RNA_dep_RNA_pol.1.
DR Pfam: PF00910; RNA_helicase.1.
DR PRINTS: PR00918; CALICVIRUSNS.
KM Polyprotein; Coat protein; Core protein; Transferase;
ET RNA-directed RNA polymerase; Hydrolase; Thiol protease.
ET CHAIN 1 27
ET CHAIN 28 249
ET CHAIN 250 495
ET CHAIN 496 795
ET CHAIN 796 984
ET CHAIN 985 1091
ET CHAIN 1092 1426
ET CHAIN 1427 1498
ET CHAIN 1499 1521
ET CHAIN 1522 1741
ET CHAIN 1742 2230
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 87.1%; Score 88; DB 1; Length 2230;
Best Local Similarity 85.0%; Pred. No. 1.8e-06;
Matches 17; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEHMKFS 20
DB 838 FSOAKISLFYTEHEHMKFS 857

RESULT 9
FDHA_METJA
ID FDHA_METJA STANDARD: PRT; 378 AA.
AC Q60314;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative formate dehydrogenase alpha chain (EC 1.2.1.2).
GN MJ0006.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
XX NCBI_Taxid=2190;
XX [1]
XX SEQUENCE FROM N.A.
RX STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE-96337999; PubMed-8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeck R., Kirness E.F., Weinstock K.G., Funtmann J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weisman J.E., Sadov P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii."
RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) = CO(2) + NADH.
CC -1- COFACTOR: MOLYBDENUM (MOLYBDOPTERIN), ZINC, FAD, MAY BIND A 4FE-4S
```

```

CC CLUSTER (BY SIMILARITY).
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC MOLYBDOPTERIN-CONTAINING
CC OXIDOREDUCTASE FAMILY.
-----
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: U67459; AAB97987.1;
CC PIR: F64300; F64300.
CC HSSP: P07658; 1AA6.
CC TIGR: M30006;
CC DR InterPro: IPR006653; Molybdop_Fe4S4.
CC DR InterPro: IPR006656; Molybdoplerin.
CC DR InterPro: IPR006655; Prok_Mdoxred.
CC Pfam: PF04879; Molybdop_Fe4S4; 1.
CC Pfam: PF00384; Molybdoplerin; 2.
CC PROSITE: PS00551; MOLYBDOPTERIN_PROK_1; 1.
CC PROSITE: PS00490; MOLYBDOPTERIN_PROK_2; 1.
CC PROSITE: PS00932; MOLYBDOPTERIN_PROK_3; FALSE_NEG.
CC DR PROSITE: PS00932; MOLYBDOPTERIN_PROK_3; FALSE_NEG.
CC KW Hypothetical protein; Oxidoreductase; Zinc; Flavoprotein; Molybdenum;
CC FAD; NAD; Iron-sulfur; 4Fe-4S; Complete proteome.
CC FT METAL 8 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC FT METAL 11 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC FT METAL 15 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC FT METAL 43 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
CC SO SEQUENCE 378 AA; 42050 MW; EDC1306DFE2609AA CRC64;
-----
Query Match 46.5%; Score 47; DB 1; Length 378;
Best Local Similarity 47.1%; Pred. No. 2.1;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
OY 4 AKISLFTTEHEIMKFS 20
DB 174 SKVITFTEHEIKLN 190
-----
RESULT 10
MARK_NYMOD
ID MARK_NYMOD STANDARD; PRT; 509 AA.
AC 098637;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
PT Maturease K (intron maturase).
MAATK.
OS Nymphaea odorata (White water lily).
OG Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Nymphaeaceae; Nymphaea.
OX NCBI_TaxID=4419;
RN [1]
RP Sequence from N.A.
RA Les D.H., Schneider E.L., Padgett D.J., Soltis P.S., Soltis D.E.,
RA Zanis M.;
RT "Phylogeny, classification and floral evolution of water lilies
RT (Nymphaeaceae; Nymphaeales): a synthesis of non-molecular, rbcL, matK
RT and 18S rDNA data.";
RL Syst. Bot. 24:28-46(1999).
CC -1- FUNCTION: Probably assists in splicing chloroplast group II
CC introns (By similarity).
CC -1- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MAATK
CC SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC CC EMBL: AF092988; AAD05556.1;
CC DR InterPro: IPR000442; Intron_maturase2.
CC DR InterPro: IPR002866; MatK_N.
CC Pfam: PF01348; Intron_maturase2; 1.
CC DR Pfam: PF01824; MatK_N; 1.
CC DR mRNA processing; Chloroplast.
CC SO SEQUENCE 509 AA; 60343 MW; C78FCA7301BDF373 CRC64;
-----
Query Match 43.6%; Score 44; DB 1; Length 509;
Best Local Similarity 72.7%; Pred. No. 9.2;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 9 FYTEHEIMKF 19
DB 472 FTEHEIVSP 482
-----
RESULT 11
PSD6_CAEEL
ID PSD6_CAEEL STANDARD; PRT; 410 AA.
AC Q20585;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable 26S proteasome non-ATPase regulatory subunit 6.
GN RPN-7 OR P49C12.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP Sequence from N.A.
RA STRAIN-Bristol N2;
RA Gardner A.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Acts as a regulatory subunit of the 26S proteasome which
CC is involved in the ATP-dependent degradation of ubiquitinated
CC proteins (By similarity).
CC -1- DOMAIN: CONTAINS 1 PCI DOMAIN.
-----
CC CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC CC EMBL: Z68227; CAA92512.1;
CC DR PIR: T22413; T22413.
CC DR Wormpep: P49C12.8; CE03368.
CC DR InterPro: IPR000717; PCI.
CC DR Pfam: PF01399; PCI; 1.
CC DR SMART: SM00088; PINT; 1.
CC KW Proteasome.
CC FT DOMAIN 205 PCI.
CC FT SEQUENCE 410 AA; 47583 MW; F37390A3250109EE CRC64;
-----
Query Match 42.6%; Score 43; DB 1; Length 410;
Best Local Similarity 36.8%; Pred. No. 11;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
OY 1 FSOAKISLFTTEHEIMKF 19
DB 156 FAMIRVGLFLDHLINKF 174
-----
RESULT 12
MARK_LATTI

```

```

ID  MATK_LATTI      STANDARD;      PRT;      506 AA.
AC  O8MCR8;
DT  28-FEB-2003 (Rel. 41, Created)
DT  28-FEB-2003 (Rel. 41, Last sequence update)
DE  28-FEB-2003 (Rel. 41, Last annotation update)
GN  Maturase K (Intron maturase).
OS  Lathyrus lingulatus (Tangier pea).
OC  Chloroplast.
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Lathyrus.
OX  NCBI_TaxID=3862;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Steele K.P., Wojciechowski M.F.;
RT  "Phylogenetic analyses of tribes Trifoliaceae and Viciaceae based on
RT  sequences of the plastid gene matk (Papilionoideae: Leguminosae).";
RT  Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
CC  -1- FUNCTION: Probably assists in splicing chloroplast group II
CC  introns (By similarity).
CC  -1- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC  SUBFAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: AF522087; AAM82079.1; -
DR  InterPro: IPR000442; Intron_maturase2.
DR  InterPro: IPR002866; Matk_N.
DR  Pfam: PF01348; Intron_maturase2; 1.
DR  Pfam: PF01824; Matk_N; 1.
DR  MRA processing: Chloroplast.
SQ  SEQUENCE 506 AA; 60897 MW; 6187B6AE73C9D564 CRC64;

Query Match      42.6%; Score 43; DB 1; Length 506;
Best Local Similarity 44.4%; Pred. No. 13;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY  2 SQAKISLFYTEHEIMK 19
DB  457 SEELQEPFTEGELISF 474

RESULT 13
YC06_KLEPN      STANDARD;      PRT;      722 AA.
AC  048452;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Putative tyrosine-protein kinase in cps region (EC 2.7.1.112) (ORF6).
OS  Klebsiella pneumoniae.
OS  Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC  Enterobacteriaceae; Klebsiella.
OX  NCBI_TaxID=573;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-Chedid;
RX  MEDLINE=95204345; PubMed=7896702;
RA  Arakawa Y., Wacharotayanankun R., Nagatsuka T., Ito H., Kato N.,
RA  Ohta M.;
RT  "Genomic organization of the Klebsiella pneumoniae cps region
RT  responsible for serotype K2 capsular polysaccharide synthesis in the
RT  virulent strain Chedid.";
RL  J. Bacteriol. 177:1788-1796(1995).
CC  -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC  tyrosine phosphate.

```

```

CC  -1- PATHWAY: Ekopolysaccharide biosynthesis.
CC  -1- SUBCELLULAR LOCATION: Inner membrane (By similarity).
CC  -1- SIMILARITY: BELONGS TO THE ETK/MZC FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL: D21242; BAA04777.1; -
DR  InterPro: IPR003856; LPS_Wzz_MPA.
DR  Pfam: PF02706; wzz; 1.
KW  Hypothetical protein; Transferase; Tyrosine-protein kinase;
KW  Ekopolysaccharide synthesis; Transmembrane; Inner membrane;
KW  ATP-binding.
FT  TRANSMEM 31 53 POTENTIAL.
FT  TRANSMEM 427 449 POTENTIAL.
SQ  SEQUENCE 722 AA; 80400 MW; 3CAD6910AEB1C3D7 CRC64;

Query Match      42.6%; Score 43; DB 1; Length 722;
Best Local Similarity 44.4%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY  1 FSOAKISLFYTEHEIMK 18
DB  326 FRESISQLYTKHEPYRK 343

RESULT 14
GCSH_LISIN      STANDARD;      PRT;      125 AA.
AC  Q928T3;
DT  28-FEB-2003 (Rel. 41, Created)
DT  28-FEB-2003 (Rel. 41, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Probable glycine cleavage system H protein.
GN  GCVH OR LIN2519.
OS  Listeria innocua.
OC  Bacteria; Firmicutes; Bacilliales; Listeriaceae; Listeria.
OX  NCBI_TaxID=1642;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-Clp 11262 / Serovar 6a;
RX  MEDLINE=21537279; PubMed=11679669;
RA  Glaser P., Franjeul L., Buchrieser C., Rusnick C., Amend A.,
RA  Baquero F., Berche P., Bloecher H., Brandt P., Chakraborty T.,
RA  Chardit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA  Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA  Entian K.-D., Esli H., Garcia-del Portillo F., Garrido P.,
RA  Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,
RA  Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkhat G.,
RA  Madueno E., Maitouran A., Mata Vicente J., Ng E., Nedjari H.,
RA  Nordlek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA  Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA  Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
RT  "Comparative genomics of Listeria species.";
RL  Science 294:849-852(2001).
CC  -1- FUNCTION: The glycine cleavage system catalyzes the degradation of
CC  glycine. The H protein shuttles the methylamine group of glycine
CC  from the P protein to the T protein (By similarity).
CC  -1- COFACTOR: The H chain contains a covalently-bound lipoyl cofactor
CC  (By similarity).
CC  -1- SUBUNIT: The glycine cleavage system is composed of four proteins:
CC  P, T, L, and H (By similarity).
CC  -1- SIMILARITY: Belongs to the gcvH family.
CC  -1- SIMILARITY: Contains 1 lipoyl-binding domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL59172; CAC9746.1; -
DR PIR: A1747; A1747.
DR Listlist: L1N02519; -
DR HAMAP: MF_00272; -; 1.
DR InterPro: IPR002930; GCV_H.
DR Pfam: PF01597; GCV_H; 1.
DR TIGRfams: TIGR00527; gcvH; 1.
DR PROSITE: PS00189; LIPOYL; 1.
DR LIPOYL: Complete proteome.
KW BINDING 63
FT SEQUENCE 125 AA; 13732 MW; 94FF0032FBCFA152 CRC64;
SQ
Query Match 41.6%; Score 42; DB 1; Length 125;
Best Local Similarity 72.7%; Pred. No. 4.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
DB 8 LFYTEHEHMK 18
      1111111111
      7 LFYTEHEHMK 17

```

```

CC -----
DR EMBL: AL591983; CAD00503.1; -
DR PIR: A11377; A11377.
DR Listlist: LMO02425; -
DR HAMAP: MF_00272; -; 1.
DR InterPro: IPR002930; GCV_H.
DR InterPro: IPR003016; LIPOYL.
DR Pfam: PF01597; GCV_H; 1.
DR TIGRfams: TIGR00527; gcvH; 1.
DR PROSITE: PS00189; LIPOYL; 1.
DR LIPOYL: Complete proteome.
KW BINDING 63
FT SEQUENCE 125 AA; 13801 MW; 1B4EE1E37C8980B5 CRC64;
SQ
Query Match 41.6%; Score 42; DB 1; Length 125;
Best Local Similarity 72.7%; Pred. No. 4.4;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
DB 8 LFYTEHEHMK 18
      1111111111
      7 LFYTEHEHMK 17

```

Search completed: October 1, 2003, 09:57:48  
 Job time : 8.58824 secs

```

RESULT 15
GCSH_LISMO STANDARD: PRT; 125 AA.
AC 08Y4L2:
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable glycine cleavage system H protein.
GN GCVH OR LMO2425.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=1167669;
RA Glaser P., Frangoul L., Buchrieser C., Rusnlok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Cherouani F., Couve E., de Daruvar A., Deboux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fajhi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstedt G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Vazquez-Boland J.-A., Voss H., Wehlant J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -I- FUNCTION: The glycine cleavage system catalyzes the degradation of
CC glycine. The H protein shuttles the methylamine group of glycine
CC from the P protein to the T protein (By similarity).
CC -I- COFACTOR: The H chain contains a covalently-bound lipoyl cofactor
CC (By similarity).
CC -I- SUBUNIT: The glycine cleavage system is composed of four proteins:
CC P, T, L, and H (By similarity).
CC -I- SIMILARITY: Belongs to the gcvH family.
CC -I- SIMILARITY: Contains 1 lipoyl-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 33.1765 Seconds

(without alignments)  
155,564 Million cell updates/sec

Title: US-09-171-432a-43

Perfect score: 101

Sequence: 1 FSQAKISLFYTEHEHMKFS 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_tvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB ID	Description
1	101	100.0	112	12	086534 hepatitis a
2	101	100.0	132	12	08V4L9 hepatitis a
3	101	100.0	132	12	08V4L6 hepatitis a
4	101	100.0	132	12	08V4M9 hepatitis a
5	101	100.0	132	12	08V4J9 hepatitis a
6	101	100.0	132	12	08V4N1 hepatitis a
7	101	100.0	132	12	08V4J5 hepatitis a
8	101	100.0	132	12	08V4M5 hepatitis a
9	101	100.0	132	12	08V4K9 hepatitis a
10	101	100.0	132	12	08V4N3 hepatitis a
11	101	100.0	132	12	08V4M7 hepatitis a
12	101	100.0	132	12	08V4N2 hepatitis a
13	101	100.0	132	12	08V4K7 hepatitis a
14	101	100.0	132	12	08V4K6 hepatitis a
15	101	100.0	132	12	08V4M4 hepatitis a
16	101	100.0	132	12	08V4N6 hepatitis a

17	101	100.0	132	12	08V4J7 hepatitis a
18	101	100.0	132	12	08V4L5 hepatitis a
19	101	100.0	132	12	08V4K2 hepatitis a
20	101	100.0	132	12	08V4K3 hepatitis a
21	101	100.0	132	12	08V4N7 hepatitis a
22	101	100.0	132	12	08V4K1 hepatitis a
23	101	100.0	132	12	08V4J8 hepatitis a
24	101	100.0	132	12	08V4L0 hepatitis a
25	101	100.0	132	12	08V4K4 hepatitis a
26	101	100.0	132	12	08V4M2 hepatitis a
27	101	100.0	132	12	08V4M3 hepatitis a
28	101	100.0	132	12	08V4L8 hepatitis a
29	101	100.0	132	12	08V4N5 hepatitis a
30	101	100.0	132	12	08V4N4 hepatitis a
31	101	100.0	132	12	08V4K8 hepatitis a
32	101	100.0	132	12	08V4M6 hepatitis a
33	101	100.0	132	12	08V4M8 hepatitis a
34	101	100.0	132	12	08V4L2 hepatitis a
35	101	100.0	132	12	08V4K0 hepatitis a
36	101	100.0	132	12	08V4L7 hepatitis a
37	101	100.0	132	12	08V4M1 hepatitis a
38	101	100.0	132	12	08V4M0 hepatitis a
39	101	100.0	132	12	08V4K5 hepatitis a
40	101	100.0	155	12	08B8M5 hepatitis a
41	101	100.0	155	12	08B8M4 hepatitis a
42	101	100.0	155	12	08B8M3 hepatitis a
43	101	100.0	155	12	08B8M2 hepatitis a
44	101	100.0	155	12	08B8M1 hepatitis a
45	101	100.0	155	12	08B8M0 hepatitis a

## ALIGNMENTS

RESULT 1

ID	086534	PRELIMINARY:	PRT:	112 AA.
AC	086534:			
DT	01-NOV-1996 (TREMBLrel. 01, Created)			
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	VP1/2A (Fragment).			
GN	VP1/2A.			
OS	Hepatitis A virus.			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;			
OC	Heptovirus.			
OX	NCBI_TaxID-12092;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-LSH/S;			
RX	MEDLINE-92348653; PubMed-1668326;			
RA	Finesch N., Cavalieri F., Garelick H., Prugnola A., Pellegrini V.,			
RA	Zuckerman A.J.;			
RT	"Characterization of a hepatitis A virus strain suitable for vaccine			
RT	production."			
RL	J. Hepatol. 13:S146-S151(1991).			
DR	EMBL: S44105; AAB22739.2; -			
FT	NON_TER			
SQ	SEQUENCE	112 AA;	13145 MW;	3COCBBB4570D9A2C CRC64;

Query Match 100.0%; Score 101; DB 12; Length 112;  
Best Local Similarity 100.0%; Pred. No. 6.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSQAKISLFYTEHEHMKFS 20  
DB 75 FSQAKISLFYTEHEHMKFS 94

RESULT 2

ID	08V4L9	PRELIMINARY:	PRT:	132 AA.
AC	08V4L9:			

DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SA-11/07/00;  
RX MEDLINE=21571641; PubMed=11714971;  
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;  
RT "Genetic analysis of hepatitis A virus strains recovered from the  
environment and from patients with acute hepatitis.";  
RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL; AF386864; AAL68525.1; -.  
FT NON\_TER 1  
FT SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;  
SQ  
QY Query Match 100.0%; Score 101; DB 12; Length 132;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 FSOAKISLFYTEHEIMKFS 20  
92 FSOAKISLFYTEHEIMKFS 111  
RESULT 3  
ID 08V4L6 PRELIMINARY; PRT; 132 AA.  
AC 08V4L6;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SA-01/09/00;  
RX MEDLINE=21571641; PubMed=11714971;  
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;  
RT "Genetic analysis of hepatitis A virus strains recovered from the  
environment and from patients with acute hepatitis.";  
RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL; AF386867; AAL68528.1; -.  
FT NON\_TER 1  
FT SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;  
SQ  
QY Query Match 100.0%; Score 101; DB 12; Length 132;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 FSOAKISLFYTEHEIMKFS 20  
92 FSOAKISLFYTEHEIMKFS 111  
RESULT 4  
ID 08V4M9 PRELIMINARY; PRT; 132 AA.  
AC 08V4M9;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SA-19/11/97;  
RX MEDLINE=21571641; PubMed=11714971;  
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;  
RT "Genetic analysis of hepatitis A virus strains recovered from the  
environment and from patients with acute hepatitis.";  
RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL; AF386854; AAL68515.1; -.  
FT NON\_TER 1  
FT SEQUENCE 132 AA; 15336 MW; D05235844D36C9D CRC64;  
SQ  
QY Query Match 100.0%; Score 101; DB 12; Length 132;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 FSOAKISLFYTEHEIMKFS 20  
92 FSOAKISLFYTEHEIMKFS 111  
RESULT 5  
ID 08V4J9 PRELIMINARY; PRT; 132 AA.  
AC 08V4J9;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-HS-06/04/99;  
RX MEDLINE=21571641; PubMed=11714971;  
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;  
RT "Genetic analysis of hepatitis A virus strains recovered from the  
environment and from patients with acute hepatitis.";  
RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL; AF386884; AAL68545.1; -.  
FT NON\_TER 1  
FT SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;  
SQ  
QY Query Match 100.0%; Score 101; DB 12; Length 132;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 FSOAKISLFYTEHEIMKFS 20  
92 FSOAKISLFYTEHEIMKFS 111  
RESULT 6  
ID 08V4N1 PRELIMINARY; PRT; 132 AA.  
AC 08V4N1;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN-SA-11/06/97;  
RA MEDLINE-21571641; PubMed-11714971;  
RX Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;  
RT "Genetic analysis of hepatitis A virus strains recovered from the  
environment and from patients with acute hepatitis.";  
RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL: AF386852; AAL68513.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 132 AA; 15354 MW; D05770901936C69D CRC64;  
  
Query Match 100.0%; Score 101; DB 12; Length 132;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 FSOAKISLFYTEHEIMKFS 20  
DB 92 FSOAKISLFYTEHEIMKFS 111  
|||||  
RESULT 7  
AC 08V4J5 PRELIMINARY; PRT; 132 AA.  
ID 08V4J5  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_Taxid=12092;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-HS-21/11/00;  
RX MEDLINE-21571641; PubMed-11714971;  
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;  
RT "Genetic analysis of hepatitis A virus strains recovered from the  
environment and from patients with acute hepatitis.";  
RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL: AF386888; AAL68549.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;  
  
Query Match 100.0%; Score 101; DB 12; Length 132;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 FSOAKISLFYTEHEIMKFS 20  
DB 92 FSOAKISLFYTEHEIMKFS 111  
|||||  
RESULT 8  
AC 08V4M5 PRELIMINARY; PRT; 132 AA.  
ID 08V4M5  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_Taxid=12092;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-SA-11/02/98;  
RX MEDLINE-21571641; PubMed-11714971;  
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;  
RT "Genetic analysis of hepatitis A virus strains recovered from the  
environment and from patients with acute hepatitis.";  
RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL: AF386850; AAL68511.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 132 AA; 15199 MW; B0C6D1A19DCF3AF2 CRC64;

RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL: AF386858; AAL68519.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 132 AA; 15284 MW; D8FC83D4BC38287D CRC64;  
  
Query Match 100.0%; Score 101; DB 12; Length 132;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 FSOAKISLFYTEHEIMKFS 20  
DB 92 FSOAKISLFYTEHEIMKFS 111  
|||||  
RESULT 9  
AC 08V4K9 PRELIMINARY; PRT; 132 AA.  
ID 08V4K9  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_Taxid=12092;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-HS-05/03/90;  
RX MEDLINE-21571641; PubMed-11714971;  
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;  
RT "Genetic analysis of hepatitis A virus strains recovered from the  
environment and from patients with acute hepatitis.";  
RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL: AF386874; AAL68535.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 132 AA; 15336 MW; D05235844D36C69D CRC64;  
  
Query Match 100.0%; Score 101; DB 12; Length 132;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 FSOAKISLFYTEHEIMKFS 20  
DB 92 FSOAKISLFYTEHEIMKFS 111  
|||||  
RESULT 10  
AC 08V4N3 PRELIMINARY; PRT; 132 AA.  
ID 08V4N3  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_Taxid=12092;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN-LI-18/12/95;  
RX MEDLINE-21571641; PubMed-11714971;  
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;  
RT "Genetic analysis of hepatitis A virus strains recovered from the  
environment and from patients with acute hepatitis.";  
RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL: AF386850; AAL68511.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 132 AA; 15199 MW; B0C6D1A19DCF3AF2 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20  
DB 92 FSOAKISLFYEEHEIMKFS 111

## RESULT 11

ID 08V4M7 PRELIMINARY; PRT; 132 AA.  
AC 08V4M7;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
NCBI\_TaxID=12092;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SA-16/01/98;  
RX MEDLINE=21571641; PubMed=11714971;  
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;  
RT "Genetic analysis of hepatitis A virus strains recovered from the  
RT environment and from patients with acute hepatitis.";  
RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL; AF386856; AAL68517.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 132 AA; 15284 MM; D8FC83D4BC38287D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20  
DB 92 FSOAKISLFYEEHEIMKFS 111

## RESULT 12

ID 08V4N2 PRELIMINARY; PRT; 132 AA.  
AC 08V4N2;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
NCBI\_TaxID=12092;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SA-12/02/96;  
RX MEDLINE=21571641; PubMed=11714971;  
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;  
RT "Genetic analysis of hepatitis A virus strains recovered from the  
RT environment and from patients with acute hepatitis.";  
RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL; AF386851; AAL68512.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 132 AA; 15268 MM; D8FC83D50C31B87D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20  
DB 92 FSOAKISLFYEEHEIMKFS 111

## RESULT 13

ID 08V4K7 PRELIMINARY; PRT; 132 AA.  
AC 08V4K7;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
NCBI\_TaxID=12092;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-HS-17/01/91;  
RX MEDLINE=21571641; PubMed=11714971;  
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;  
RT "Genetic analysis of hepatitis A virus strains recovered from the  
RT environment and from patients with acute hepatitis.";  
RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL; AF386876; AAL68537.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 132 AA; 15336 MM; D05235844D36C69D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20  
DB 92 FSOAKISLFYEEHEIMKFS 111

## RESULT 14

ID 08V4K6 PRELIMINARY; PRT; 132 AA.  
AC 08V4K6;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
NCBI\_TaxID=12092;

RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-HS-27/02/91;  
RX MEDLINE=21571641; PubMed=11714971;  
RA Pina S., Buti M., Jardi R., Clemente-Casares P., Jofre J., Girones R.;  
RT "Genetic analysis of hepatitis A virus strains recovered from the  
RT environment and from patients with acute hepatitis.";  
RL J. Gen. Virol. 82:2955-2963(2001).  
DR EMBL; AF386877; AAL68538.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 132 AA; 15266 MM; D8FC83D4B97D3C29 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20  
DB 92 FSOAKISLFYEEHEIMKFS 111



## RESULT 15

Q8V4M4 PRELIMINARY; PRT; 132 AA.  
 AC Q8V4M4;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 OS Hepatitis A virus.  
 OS Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Hepatovirus.  
 OC NCBI\_TaxID=12092;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SA-22/03/99;  
 RX MEDLINE=21571641; PubMed=11714971;  
 RA Pina S., Buti M., Jordi R., Clemente-Casares P., Jofre J., Girones R.,  
 RT "Genetic analysis of hepatitis A virus strains recovered from the  
 RT environment and from patients with acute hepatitis.";  
 RT J. Gen. Virol. 82:2955-2963(2001).  
 EMBL; AF386859; AAL68520.1; -.  
 NON\_TER 1  
 FT 132 132  
 SQ SEQUENCE 132 AA; 15256 MW; B06D0ACC35383AD8 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLPTTEHEIMKFS 20  
 ||||||||||||||||  
 DB 92 FSOAKISLPTTEHEIMKFS 111

Search completed: October 1, 2003, 10:02:38  
 Job time : 35.1765 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 ; Search time 12.7059 Seconds  
(without alignments)  
66.600 Million cell updates/sec

Title: US-09-171-432a-43

Perfect score: 101

Sequence: 1 FSOAKISLFTYEHEIMKFS 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310658 residues

al number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2.6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2.6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2.6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2.6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2.6/ptodata/1/1aa/PCITUS.COMB.pep:\*  
6: /cgn2.6/ptodata/1/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	1091	6	5516630-2
2	101	100.0	2227	3	US-08-475-886-2
3	101	100.0	2227	3	US-08-475-886-4
4	101	100.0	2227	3	US-08-475-886-6
5	101	100.0	2227	3	US-08-397-232-2
6	101	100.0	2227	3	US-08-397-232-4
7	101	100.0	2227	3	US-09-171-387-2
8	101	100.0	2227	4	US-09-653-499-2
9	101	100.0	2227	4	US-09-653-499-4
10	101	100.0	2227	4	US-09-653-499-6
11	43	40.6	171	4	US-09-134-001C-5205
12	41	40.6	165	4	US-09-107-532A-7030
13	41	40.6	173	4	US-09-585-228-4
14	41	40.6	2431	1	US-07-920-281C-2
15	41	40.6	2431	3	US-08-466-277-2
16	40	39.6	173	4	US-09-585-228-2
17	40	39.6	240	4	US-09-328-352-7537
18	40	39.6	243	4	US-09-252-991A-17307
19	39	38.6	148	4	US-09-198-452A-491
20	39	38.6	170	4	US-09-198-119C-89
21	39	38.6	289	4	US-09-305-856B-8
22	39	38.6	289	5	PCT-US92-00282-11
23	39	38.6	340	4	US-09-107-532A-3720
24	39	38.6	455	4	US-09-198-452A-372
25	39	38.6	530	4	US-09-252-991A-23666
26	39	38.6	668	4	US-09-252-991A-21541
27	39	38.6	860	1	US-08-092-817-4

28	39	38.6	860	4	US-08-485-128-4	Sequence 4, Appl1
29	39	38.6	860	4	US-09-804-778A-8	Sequence 8, Appl1
30	39	38.6	1410	2	US-08-470-058-4	Sequence 4, Appl1
31	39	38.6	1410	3	US-09-037-188-4	Sequence 4, Appl1
32	39	38.6	1410	3	US-09-285-310-4	Sequence 4, Appl1
33	38	37.6	265	1	US-07-958-551-2	Sequence 2, Appl1
34	38	37.6	265	1	US-08-129-610-7	Sequence 7, Appl1
35	38	37.6	265	1	US-08-129-609A-7	Sequence 7, Appl1
36	38	37.6	265	1	US-08-455-313-7	Sequence 7, Appl1
37	38	37.6	265	1	US-08-475-924-2	Sequence 2, Appl1
38	38	37.6	265	2	US-08-657-579A-2	Sequence 2, Appl1
39	38	37.6	265	3	US-09-224-023-7	Sequence 7, Appl1
40	38	37.6	265	5	PCT-US94-07887-7	Sequence 7, Appl1
41	38	37.6	390	4	US-09-328-352-7807	Sequence 7807, Ap
42	38	37.6	707	3	US-08-772-270A-4	Sequence 4, Appl1
43	38	37.6	707	4	US-09-062-126-4	Sequence 4, Appl1
44	37.5	37.1	86	4	US-09-328-352-6252	Sequence 6252, Ap
45	37.5	37.1	378	4	US-09-107-532A-6500	Sequence 6500, Ap

## ALIGNMENTS

```
RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; FEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANIELLO, VINCENT R.;
; BAROUDY, BAIRGE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO: 2:
; LENGTH: 1091
5516630-2

Query Match      100.0%; Score 101; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 1e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSOAKISLFTYEHEIMKFS 20
DB      1071 FSOAKISLFTYEHEIMKFS 1090

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SOZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262052
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
```

SEQ ID NO 2  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-08-475-886-2

Query Match 100.0%; Score 101; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20  
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 3  
US-08-475-886-4  
Sequence 4, Application US/08475886A  
Patent No. 6113912  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 202642620S2  
CURRENT APPLICATION NUMBER: US/08/475,886A  
CURRENT FILING DATE: 1995-06-07  
EARLIER APPLICATION NUMBER: 07/947,338  
EARLIER FILING DATE: 1992-09-18  
EARLIER APPLICATION NUMBER: 08/397,232  
EARLIER FILING DATE: 1993-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
US-08-475-886-4

Query Match 100.0%; Score 101; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20  
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 4  
US-08-475-886-6  
Sequence 6, Application US/08475886A  
Patent No. 6113912  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 202642620S2  
CURRENT APPLICATION NUMBER: US/08/475,886A  
CURRENT FILING DATE: 1995-06-07  
EARLIER APPLICATION NUMBER: 07/947,338  
EARLIER FILING DATE: 1992-09-18  
EARLIER APPLICATION NUMBER: 08/397,232  
EARLIER FILING DATE: 1993-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 6  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-08-475-886-6

Query Match 100.0%; Score 101; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20  
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 5  
US-08-397-232-2  
Sequence 2, Application US/08397232A  
Patent No. 6180110  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 202642620S1  
CURRENT APPLICATION NUMBER: US/08/397,232A  
CURRENT FILING DATE: 1995-04-17  
EARLIER APPLICATION NUMBER: 07/947,338  
EARLIER FILING DATE: 1992-09-18  
EARLIER APPLICATION NUMBER: PCT/US93/08610  
EARLIER FILING DATE: 1993-09-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 2  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-08-397-232-2

Query Match 100.0%; Score 101; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20  
Db 834 FSOAKISLFYTEHEIMKFS 853

RESULT 6  
US-08-397-232-4  
Sequence 4, Application US/08397232A  
Patent No. 6180110  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 202642620S1  
CURRENT APPLICATION NUMBER: US/08/397,232A  
CURRENT FILING DATE: 1995-04-17  
EARLIER APPLICATION NUMBER: 07/947,338  
EARLIER FILING DATE: 1992-09-18  
EARLIER APPLICATION NUMBER: PCT/US93/08610  
EARLIER FILING DATE: 1993-09-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-08-397-232-4

Query Match 100.0%; Score 101; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFTTEHEIMKFS 20  
Db 834 FSOAKISLFTTEHEIMKFS 853

## RESULT 7

US-09-171-387-2  
Sequence 2, Application US/09171387  
Patent No. 6280734

## GENERAL INFORMATION:

APPLICANT: RAYCHAUDHURI, GOPA;  
EMERSON, SUZANNE, U.;

PURCELL, ROBERT, H.;

TITLE OF INVENTION: SIMIAN-HUMAN HAV  
HAVING A CHIMERIC 2C PROTEIN

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESSER: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: IBM PC COMPATIBLE

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: MICROSOFT WORD 97

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/171,387

FILING DATE: 24-Mar-1999

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US97/06506

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US60/015,642

FILING DATE: 19-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: William S. Feiler

REGISTRATION NUMBER: 26,728

REFERENCE/DOCKET NUMBER: 2026-4229US1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 758-4800

TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 2

SEQUENCE CHARACTERISTICS:

LENGTH: 2227 amino acids

TYPE: amino acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query Match 100.0%; Score 101; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFTTEHEIMKFS 20  
Db 834 FSOAKISLFTTEHEIMKFS 853

## RESULT 8

US-09-653-499-2

Sequence 2, Application US/09653499

Patent No. 6423318

GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W

APPLICANT: EMERSON, SUZANNE U

APPLICANT: PURCELL, ROBERT H

APPLICANT: D'HONDT, ERIC

TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

FILE REFERENCE: 20264262US2

CURRENT APPLICATION NUMBER: US/09/653,499

CURRENT FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: 08/475,886  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 2

LENGTH: 2227

TYPE: PRT

ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175

US-09-653-499-2

Query Match 100.0%; Score 101; DB 4; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFTTEHEIMKFS 20  
Db 834 FSOAKISLFTTEHEIMKFS 853

## RESULT 9

US-09-653-499-4

Sequence 4, Application US/09653499

Patent No. 6423318

GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W

APPLICANT: EMERSON, SUZANNE U

APPLICANT: PURCELL, ROBERT H

APPLICANT: D'HONDT, ERIC

TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

FILE REFERENCE: 20264262US2

CURRENT APPLICATION NUMBER: US/09/653,499

CURRENT FILING DATE: 2000-08-31

PRIOR APPLICATION NUMBER: 08/475,886

PRIOR FILING DATE: 1999-08-09

PRIOR APPLICATION NUMBER: 07/947,338

PRIOR FILING DATE: 1992-09-18

PRIOR APPLICATION NUMBER: 08/397,232

PRIOR FILING DATE: 1995-03-10

NUMBER OF SEQ ID NOS: 6

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 4

LENGTH: 2227

TYPE: PRT

ORGANISM: Attenuated HAV (Pass 35), strain HM-175

US-09-653-499-4

Query Match 100.0%; Score 101; DB 4; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.3e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFTTEHEIMKFS 20  
Db 834 FSOAKISLFTTEHEIMKFS 853

## RESULT 10

US-09-653-499-6

Sequence 6, Application US/09653499

Patent No. 6423318

GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W

APPLICANT: EMERSON, SUZANNE U

APPLICANT: PURCELL, ROBERT H

APPLICANT: D'HONDT, ERIC

TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

FILE REFERENCE: 20264262US2

CURRENT APPLICATION NUMBER: US/09/653,499

US-09-653-499-6

```

: PRIOR APPLICATION NUMBER: 08/475,886
: PRIOR FILING DATE: 1999-08-09
: PRIOR APPLICATION NUMBER: 07/947,338
: PRIOR FILING DATE: 1992-09-18
: PRIOR APPLICATION NUMBER: 08/397,232
: PRIOR FILING DATE: 1995-03-10
: NUMBER OF SEQ ID NOS: 6
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 6
: LENGTH: 2227
: TYPE: PRF
: ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match
Best Local Similarity 100.0%; Score 101; DB 4; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20
      |||
      834 FSOAKISLFYEEHEIMKFS 853

RESULT 11
US-09-134-001C-5205
: Sequence 5205, Application US/09134001C
: Patent No. 6380370
: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
: FILE REFERENCE: CTC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: PRIOR FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 5205
: LENGTH: 171
: TYPE: PRF
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5205

Query Match
Best Local Similarity 42.6%; Score 43; DB 4; Length 171;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 SOAKISLFYEEHE 15
      |||
      93 SOAKISLFYEEHE 106

DB

RESULT 12
US-09-107-532A-7030
: Sequence 7030, Application US/09107532A
: Patent No. 6583275
: GENERAL INFORMATION:
: APPLICANT: Lynn A Doucette-Stamm and David Bush
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
: NUMBER OF SEQUENCES: 7310
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GENOME THERAPEUTICS CORPORATION
: STREET: 100 Beaver Street
: CITY: Waltham
: STATE: Massachusetts
: COUNTRY: USA
: ZIP: 02354
: COMPUTER READABLE FORM:
: MEDIUM TYPE: CD-ROM ISO9660
: COMPUTER: PC
```

```

: OPERATING SYSTEM: <Unknown>
: SOFTWARE: ASCII
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/107,532A
: FILING DATE: 30-Jun-1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 60/085,598
: FILING DATE: 14 May 1998
: APPLICATION NUMBER: 60/051571
: FILING DATE: July 2, 1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Arinello, Pamela Deneke
: REGISTRATION NUMBER: 40,489
: REFERENCE/DOCKET NUMBER: GTC-012
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (781)893-5007
: TELEFAX: (781)893-8277
: INFORMATION FOR SEQ ID NO: 7030:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 165 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: HYPOTHEICAL: YES
: ORIGINAL SOURCE:
: ORGANISM: Enterococcus faecium
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (B) LOCATION 1...165
: SEQUENCE DESCRIPTION: SEQ ID NO: 7030:
US-09-107-532A-7030

Query Match
Best Local Similarity 40.6%; Score 41; DB 4; Length 165;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 5 KISLFYEEHEIMK 18
      |||
      1 KISYIYIRERVMK 14

DB

RESULT 13
US-09-585-228-4
: Sequence 4, Application US/09585228
: Patent No. 6531576
: GENERAL INFORMATION:
: APPLICANT: Piddington, Christopher S.
: APPLICANT: West, James W.
: APPLICANT: Holly, Richard D.
: APPLICANT: Burkhead, Steven K.
: TITLE OF INVENTION: FOUR-HELICAL BUNDLE PROTEIN ZSIG81
: FILE REFERENCE: 99-13
: CURRENT APPLICATION NUMBER: US/09/585,228
: CURRENT FILING DATE: 2000-06-01
: EARLIER APPLICATION NUMBER: US 60/137,057
: EARLIER FILING DATE: 1999-06-01
: NUMBER OF SEQ ID NOS: 36
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 4
: LENGTH: 173
: TYPE: PRF
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: SIGNAL
: LOCATION: (1)...(17)
US-09-585-228-4

Query Match
Best Local Similarity 40.6%; Score 41; DB 4; Length 173;
Matches 8; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 FSOAKISLFYEEHEIMKFS 20
      |||
      1 FSOAKISLFYEEHEIMKFS 20
```

Db 148 WAQSLISLFQALRHDLMSS 167

## RESULT 14

US-07-920-281C-2  
Sequence 2, Application US/07920281C

Patent No. 5739026

GENERAL INFORMATION:

APPLICANT: Garoff, Henrik

APPLICANT: Liljestrom, Peter

TITLE OF INVENTION: DNA Expression Systems Based on

TITLE OF INVENTION: Alphaviruses

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch &amp; Birch

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/920,281C

FILING DATE: 13-AUG-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M.

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 828-103P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-241-1300

TELEFAX: 703-241-2848

TELEX: 248345

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2431 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-07-920-281C-2

Query Match 40.6%; Score 41; DB 1; Length 2431;

Best Local Similarity 50.0%; Pred. No. 2.9e+02;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 1072 FSAPKVSLEYNNH 1085

OY 1 FSQAKISLFYTEEH 14

US-08-466-277-2

Sequence 2, Application US/08466277

Patent No. 6190666

GENERAL INFORMATION:

APPLICANT: Garoff, Henrik

APPLICANT: Liljestrom, Peter

TITLE OF INVENTION: DNA Expression Systems Based on

TITLE OF INVENTION: Alphaviruses

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch &amp; Birch

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/466,277

FILING DATE: 06-Jun-1995

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/920,281

FILING DATE: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: Murphy Jr., Gerald M.

REGISTRATION NUMBER: 28,977

REFERENCE/DOCKET NUMBER: 828-103P

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-241-1300

TELEFAX: 703-241-2848

TELEX: 248345

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2431 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-08-466-277-2

Query Match 40.6%; Score 41; DB 3; Length 2431;

Best Local Similarity 50.0%; Pred. No. 2.9e+02;

Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Db 1072 FSAPKVSLEYNNH 1085

OY 1 FSQAKISLFYTEEH 14

US-08-466-277-2

Sequence 2, Application US/08466277

Patent No. 6190666

GENERAL INFORMATION:

APPLICANT: Garoff, Henrik

APPLICANT: Liljestrom, Peter

TITLE OF INVENTION: DNA Expression Systems Based on

TITLE OF INVENTION: Alphaviruses

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Birch, Stewart, Kolasch &amp; Birch

STREET: P.O. Box 747

CITY: Falls Church

STATE: Virginia

COUNTRY: USA

ZIP: 22040-0747

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

Search completed: October 1, 2003, 10:06:30  
Job time: 13.7059 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: October 1, 2003, 10:02:52 ; Search time 24.1176 Seconds  
(without alignments)  
131.201 Million cell updates/sec

Title: US-09-171-432a-43  
Perfect score: 101  
Sequence: 1 FSOAKISLFYTEHEIMKFS 20

Scoring table:  
BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 587654 seqs, 158212981 residues  
Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10C\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	101	100.0	352	US-10-272-459-45	Sequence 45, Appl
2	101	100.0	980	US-10-272-459-41	Sequence 41, Appl
3	101	100.0	2227	US-09-928-955-12	Sequence 12, Appl
4	101	100.0	2227	US-10-104-966-12	Sequence 12, Appl
5	101	100.0	2227	US-10-135-988-2	Sequence 2, Appl
6	101	100.0	2227	US-10-135-988-4	Sequence 4, Appl
7	101	100.0	2227	US-10-135-988-6	Sequence 6, Appl
8	47	46.5	518	US-09-253-088-6	Sequence 6, Appl
9	47	46.5	816	US-09-253-088-5	Sequence 5, Appl
10	41	40.6	1336	US-10-224-999A-3482	Sequence 3482, Ap
11	41	40.6	2431	US-09-901-106-2	Sequence 2, Appl
12	40	39.6	173	US-10-199-672-506	Sequence 506, App
13	40	39.6	173	US-10-187-749-506	Sequence 506, App
14	40	39.6	173	US-10-194-457-506	Sequence 506, App
15	40	39.6	173	US-10-184-642-506	Sequence 506, App

16	40	39.6	173	12	US-10-196-747-506	Sequence 506, App
17	40	39.6	173	12	US-10-173-689-506	Sequence 506, App
18	40	39.6	173	12	US-10-173-690-506	Sequence 506, App
19	40	39.6	173	12	US-10-173-691-506	Sequence 506, App
20	40	39.6	173	12	US-10-173-692-506	Sequence 506, App
21	40	39.6	173	12	US-10-173-693-506	Sequence 506, App
22	40	39.6	173	12	US-10-173-694-506	Sequence 506, App
23	40	39.6	173	12	US-10-173-695-506	Sequence 506, App
24	40	39.6	173	12	US-10-173-707-506	Sequence 506, App
25	40	39.6	173	12	US-10-173-707-506	Sequence 506, App
26	40	39.6	173	12	US-10-174-583-506	Sequence 506, App
27	40	39.6	173	12	US-10-174-587-506	Sequence 506, App
28	40	39.6	173	12	US-10-174-589-506	Sequence 506, App
29	40	39.6	173	12	US-10-174-591-506	Sequence 506, App
30	40	39.6	173	12	US-10-175-736-506	Sequence 506, App
31	40	39.6	173	12	US-10-175-742-506	Sequence 506, App
32	40	39.6	173	12	US-10-175-744-506	Sequence 506, App
33	40	39.6	173	12	US-10-175-745-506	Sequence 506, App
34	40	39.6	173	12	US-10-175-748-506	Sequence 506, App
35	40	39.6	173	12	US-10-175-751-506	Sequence 506, App
36	40	39.6	173	12	US-10-175-754-506	Sequence 506, App
37	40	39.6	173	12	US-10-176-480-506	Sequence 506, App
38	40	39.6	173	12	US-10-176-489-506	Sequence 506, App
39	40	39.6	173	12	US-10-176-754-506	Sequence 506, App
40	40	39.6	173	12	US-10-176-755-506	Sequence 506, App
41	40	39.6	173	12	US-10-176-759-506	Sequence 506, App
42	40	39.6	173	12	US-10-176-920-506	Sequence 506, App
43	40	39.6	173	12	US-10-176-922-506	Sequence 506, App
44	40	39.6	173	12	US-10-176-924-506	Sequence 506, App
45	40	39.6	173	12	US-10-176-984-506	Sequence 506, App

## ALIGNMENTS

RESULT 1  
US-10-272-459-45  
; Sequence 45, Application US/10272459  
; Publication No. US20030124517A1  
; GENERAL INFORMATION:  
; APPLICANT: PICHONANTS, Sergio  
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT  
; FILE OF INVENTION: PROTEINS AND USES THEREOF  
; FILE REFERENCE: P11955.002 / 2301-17955  
; CURRENT APPLICATION NUMBER: US/10/272,459  
; CURRENT FILING DATE: 2002-10-15  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein  
; OTHER INFORMATION: of 38.8 kDa  
US-10-272-459-45

Query Match 100.0%; Score 101; DB 15; Length 352;  
Best Local Similarity 100.0%; Pred No. 1.3e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 FSOAKISLFYTEHEIMKFS 20  
DB 206 FSOAKISLFYTEHEIMKFS 225  
RESULT 2  
US-10-272-459-41  
; Sequence 41, Application US/10272459  
; Publication No. US20030124517A1  
; GENERAL INFORMATION:  
; APPLICANT: PICHONANTS, Sergio  
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

```
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41

Query Match          100.0%; Score 101; DB 15; Length 980;
Best Local Similarity 100.0%; Pred. No. 4e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSOAKISLFYTEEHIMKFS 20
        |||
        834 FSOAKISLFYTEEHIMKFS 853

RESULT 3
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
09-929-955-12

Query Match          100.0%; Score 101; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSOAKISLFYTEEHIMKFS 20
        |||
        834 FSOAKISLFYTEEHIMKFS 853

RESULT 4
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
```

```
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match          100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSOAKISLFYTEEHIMKFS 20
        |||
        834 FSOAKISLFYTEEHIMKFS 853

RESULT 5
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262U3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match          100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 9.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 FSOAKISLFYTEEHIMKFS 20
        |||
        834 FSOAKISLFYTEEHIMKFS 853

RESULT 6
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 202642620S3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
```

```

; PRT
; ORGANISM: group B streptococcus
; US-09-252-088-6
Query Match
Best Local Similarity 100.0%; Score 101; DB 14; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 834 FSOAKISLFYTEHEIMKFS 853

RESULT 7
US-10-135-988-6
Sequence 6, Application US/10135988
Publication No. US20020176869A1
GENERAL INFORMATION:
APPLICANT: PUNKHUSER, ANN M
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 2026426203
CURRENT APPLICATION NUMBER: US/10/135,988
CURRENT FILING DATE: 2002-04-29
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1993-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin Ver. 2.1
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-10-135-988-6
Query Match
Best Local Similarity 100.0%; Score 101; DB 14; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 834 FSOAKISLFYTEHEIMKFS 853

RESULT 8
US-09-252-088-6
Sequence 6, Application US/09252088
Publication No. US20030031682A1
GENERAL INFORMATION:
APPLICANT: BRODEUR, Bernard R.
APPLICANT: RIOUX, Clement
APPLICANT: BOYER, Martine
APPLICANT: CHARLEBOIS, Isabelle
APPLICANT: HAMEL, Jose
TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS
FILE REFERENCE: 8331-9002
CURRENT APPLICATION NUMBER: US/09/252,088
CURRENT FILING DATE: 1999-02-18
EARLIER APPLICATION NUMBER: US/60/075,425
EARLIER FILING DATE: 1998-02-20
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 6
LENGTH: 518
```

```

; PRT
; ORGANISM: group B streptococcus
; US-09-252-088-6
Query Match
Best Local Similarity 46.5%; Score 47; DB 11; Length 518;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 371 FOQAPALVKNAEHYVAF 390

RESULT 9
US-09-252-088-5
Sequence 5, Application US/09252088
Publication No. US20030031682A1
GENERAL INFORMATION:
APPLICANT: BRODEUR, Bernard R.
APPLICANT: RIOUX, Clement
APPLICANT: BOYER, Martine
APPLICANT: CHARLEBOIS, Isabelle
APPLICANT: HAMEL, Jose
APPLICANT: MARTIN, Denis
TITLE OF INVENTION: NOVEL GROUP B STREPTOCOCCUS ANTIGENS
FILE REFERENCE: 8331-9002
CURRENT APPLICATION NUMBER: US/09/252,088
CURRENT FILING DATE: 1999-02-18
EARLIER APPLICATION NUMBER: US/60/075,425
EARLIER FILING DATE: 1998-02-20
NUMBER OF SEQ ID NOS: 44
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 816
TYPE: PRT
ORGANISM: group B streptococcus
US-09-252-088-5
Query Match
Best Local Similarity 46.5%; Score 47; DB 11; Length 816;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEHEIMKFS 20
DB 669 FOQAPALVKNAEHYVAF 688

RESULT 10
US-10-224-999A-3482
Sequence 3482, Application US/10224999A
Publication No. US20030171318A1
GENERAL INFORMATION:
APPLICANT: Myriad Genetics, Inc.
APPLICANT: Morham, Scott
APPLICANT: Zavitz, Kenton
APPLICANT: Hobden, Adrian
TITLE OF INVENTION: Composition and Method for Treating Viral Infection
FILE REFERENCE: 5004.01
CURRENT APPLICATION NUMBER: US/10/224,999A
CURRENT FILING DATE: 2003-03-03
PRIOR APPLICATION NUMBER: US 60/313,695
PRIOR FILING DATE: 2001-08-20
NUMBER OF SEQ ID NOS: 3484
SOFTWARE: Patentin version 3.1
SEQ ID NO 3482
LENGTH: 1336
TYPE: PRT
ORGANISM: Semliki Forest virus
US-10-224-999A-3482
Query Match
Best Local Similarity 40.6%; Score 41; DB 12; Length 1336;
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

OY 1 FSOAKISLFYTEEH 14  
DB 1072 FSAKPSLVYENH 1085

## RESULT 11

US-09-901-106-2  
Sequence 2, Application US/09901106  
Patent No. US20020151067A1  
GENERAL INFORMATION:  
APPLICANT: Garoff, Henrik  
Liljestrom, Peter  
TITLE OF INVENTION: DNA Expression Systems Based on  
Alphaviruses  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Birch, Stewart, Kolasch & Birch  
STREET: P.O. Box 747  
CITY: Falls Church  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22040-0747  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/901.106  
FILING DATE: 10-Jul-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/920.281C  
FILING DATE: 13-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Murphy Jr., Gerald M.  
REGISTRATION NUMBER: 28.977  
REFERENCE/DOCKET NUMBER: 828-103P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-241-1300  
TELEFAX: 703-241-2848  
TELEX: 248345  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2431 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
09-901-106-2

Query Match 40.6%; Score 41; DB 10; Length 2431;  
Best Local Similarity 50.0%; Pred. No. 7.7e+02;  
Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEEH 14  
DB 1072 FSAKPSLVYENH 1085

## RESULT 12

US-10-199-672-506  
Sequence 506, Application US/10199672  
Publication No. US20030148442A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
Chen, Jian  
APPLICANT: Desnoyers, Luc  
Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
Gurney, Austin L.  
APPLICANT: Pan, James

APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
ACIDS ENCODING THE SAME  
FILE REFERENCE: P3430R1C1  
CURRENT APPLICATION NUMBER: US/10/199.672  
CURRENT FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: US/10/052.586  
PRIOR FILING DATE: 2002-01-15  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059266  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/063120  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063121  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063486  
PRIOR FILING DATE: 1997-10-21  
PRIOR APPLICATION NUMBER: 60/063540  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063541  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063544  
PRIOR FILING DATE: 1997-10-28  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 612  
SEQ ID NO 506  
LENGTH: 173  
TYPE: PRT  
ORGANISM: Homo Sapien

US-10-199-672-506

Query Match 39.6%; Score 40; DB 12; Length 173;  
Best Local Similarity 35.0%; Pred. No. 61;  
Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

OY 1 FSOAKISLFYTEEHIMKFS 20  
DB 148 WAOSLVSLFQALRHDLMRSS 167

## RESULT 13

US-10-187-749-506  
Sequence 506, Application US/10187749  
Publication No. US20030135036A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
Chen, Jian  
APPLICANT: Desnoyers, Luc  
Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
Gurney, Austin L.  
APPLICANT: Pan, James  
APPLICANT: Smith, Victoria  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
ACIDS ENCODING THE SAME  
FILE REFERENCE: P3430R1C1  
CURRENT APPLICATION NUMBER: US/10/187.749  
CURRENT FILING DATE: 2002-07-01  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059266  
PRIOR FILING DATE: 1997-09-18

```
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063120
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063121
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063486
;; PRIOR FILING DATE: 1997-10-21
;; PRIOR APPLICATION NUMBER: 60/063540
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063541
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063544
;; PRIOR FILING DATE: 1997-10-28
;; Remaining Prior Application data removed - See file wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 506
;; LENGTH: 173
;; TYPE: PRT
;; ORGANISM: Homo Sapien
10-187-749-506

Query Match
Best Local Similarity 39.6%; Score 40; DB 12; Length 173;
Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
DB 148 WAQSLVSLFQALRHDLRRSS 167

RESULT 14
US-10-194-457-506
;; Sequence 506, Application US/10194457
;; Publication No. US20030153037A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Chen, Jian
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Pan, James
;; APPLICANT: Smith, Victoria
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; TITLE OF INVENTION: ACIDS ENCODING THE SAME
;; FILE REFERENCE: P3430RIC296
;; CURRENT APPLICATION NUMBER: US/10/194,457
;; CURRENT FILING DATE: 2002-07-11
;; PRIOR APPLICATION NUMBER: 10/052586
;; PRIOR FILING DATE: 2002-01-15
;; PRIOR APPLICATION NUMBER: 60/059263
;; PRIOR FILING DATE: 1997-09-18
;; PRIOR APPLICATION NUMBER: 60/059266
;; PRIOR FILING DATE: 1997-09-18
;; PRIOR APPLICATION NUMBER: 60/062250
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063120
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063121
;; PRIOR FILING DATE: 1997-10-24
;; PRIOR APPLICATION NUMBER: 60/063486
;; PRIOR FILING DATE: 1997-10-21
;; PRIOR APPLICATION NUMBER: 60/063540
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063541
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/063544
;; PRIOR FILING DATE: 1997-10-28
;; Prior Application data removed - See file wrapper or PALM.
```

```
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 506
;; LENGTH: 173
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-194-457-506

Query Match
Best Local Similarity 39.6%; Score 40; DB 12; Length 173;
Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
DB 148 WAQSLVSLFQALRHDLRRSS 167

RESULT 15
US-10-184-642-506
;; Sequence 506, Application US/10184642
;; Publication No. US20030157635A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin P.
;; APPLICANT: Chen, Jian
;; APPLICANT: Desnoyers, Luc
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Pan, James
;; APPLICANT: Smith, Victoria
;; APPLICANT: Watanabe, Colin K.
;; APPLICANT: Wood, William I.
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; TITLE OF INVENTION: ACIDS ENCODING THE SAME
;; FILE REFERENCE: P3430RIC194
;; CURRENT APPLICATION NUMBER: US/10/184,642
;; CURRENT FILING DATE: 2002-06-27
;; PRIOR APPLICATION removed - See file wrapper or Palm
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 506
;; LENGTH: 173
;; TYPE: PRT
;; ORGANISM: Homo Sapien
US-10-184-642-506

Query Match
Best Local Similarity 39.6%; Score 40; DB 12; Length 173;
Matches 7; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FSOAKISLFYTEHEIMKFS 20
DB 148 WAQSLVSLFQALRHDLRRSS 167
```

Search completed: October 1, 2003, 10:37:51  
Job time : 25.1176 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 ; Search time 43.5294 Seconds  
(without alignments)  
72.928 Million cell updates/sec

Title: US-09-171-432A-44  
Perfect score: 104  
Sequence: 1 KVNPRHGMLEIRANSKD 20

Scoring table: BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1107863 seqs, 158726573 residues  
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.GeneSeq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneSeq/geneSeq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	104	100.0	20	18 AAM42927
2	104	100.0	21	22 AAB6444
3	104	100.0	2227	7 AAP6066
4	104	100.0	2227	11 AAR0597
5	104	100.0	2227	18 AAM34074
6	104	100.0	2227	21 AAB18607
7	104	100.0	2227	21 AAB18608
8	104	100.0	2227	21 AAB18609
9	104	100.0	2227	23 ABG31727

10	104	100.0	2227	23 ABG31728	Hepatitis A virus
11	104	100.0	2227	23 ABG31729	Attenuated Hepatit
12	104	100.0	2227	23 ABB19899	Hepatitis A virus
13	104	100.0	2227	24 ABU08639	Wild type human he
14	104	100.0	2227	24 ABU08640	Attenuated (pass3
15	104	100.0	2227	24 ABU08641	Immunogenic Hepati
16	53	51.0	20	18 AAM42928	Synthetic HAV P2A
17	53	51.0	21	22 AAB69445	Lactococcus lactis
18	47	45.2	387	23 ABB53661	Novel human secret
19	45	43.3	195	22 AAU32447	Arbidopais thalia
20	45	43.3	206	21 AAG32341	Arbidopais thalia
21	45	43.3	242	21 AAG10097	Arbidopais thalia
22	45	43.3	242	21 AAG32340	Arbidopais thalia
23	45	43.3	383	21 AAG10096	Arbidopais thalia
24	45	43.3	383	21 AAG32339	Arbidopais thalia
25	45	43.3	399	21 AAG10095	Arbidopais thalia
26	44	42.3	102	22 ABG11781	Novel human diagno
27	44	42.3	432	20 AAY49167	Human succinyl-Coe
28	44	42.3	432	21 AAY76886	Human SCSH-2 prote
29	44	42.3	432	22 AAU32446	Novel human secret
30	43.5	41.8	267	21 AAG37721	Arbidopais thalia
31	43.5	41.8	267	21 AAG37736	Arbidopais thalia
32	43.5	41.8	269	21 AAG24210	Arbidopais thalia
33	43.5	41.8	426	21 AAG37720	Arbidopais thalia
34	43.5	41.8	426	21 AAG37735	Arbidopais thalia
35	43.5	41.8	426	23 ABB93930	Herbicidally activ
36	43.5	41.8	428	21 AAG24209	Arbidopais thalia
37	43.5	41.8	438	21 AAG37734	Arbidopais thalia
38	43.5	41.8	440	21 AAG24208	Arbidopais thalia
39	43.5	41.8	442	21 AAG37719	Arbidopais thalia
40	43	41.3	98	20 AAY29645	Human secreted pro
41	43	41.3	98	20 AAY39330	Protein product of
42	43	41.3	98	21 AAB44333	Human PRO725 prote
43	43	41.3	98	22 AAB44893	Human angiotensin
44	43	41.3	98	23 ABB95469	Human PRO725 prote
45	43	41.3	98	23 ABB84863	Human PRO725 prote

## ALIGNMENTS

RESULT 1	
AAW42927	standard; peptide; 20 AA.
ID	AAW42927
XX	
AC	AAW42927;
XX	
DT	28-APR-1998 (first entry)
XX	
DE	Immunogenic Hepatitis A virus peptide YK-1327.
XX	
KW	Immunogenic peptide; immunogenic epitope; P2A protein;
KM	Immune response; antibody.
XX	
OS	Synthetic.
OS	Hepatitis A virus.
XX	
PN	WO9740147-A1.
XX	
PD	30-OCT-1997.
XX	
XX	
PF	18-APR-1997; 97WO-US06891.
XX	
PR	19-APR-1996; 96US-0015644.
XX	
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX	
PI	Fields HA, Khudyakov YE.
XX	
DR	WPI; 1997-535831/49.
XX	
PT	Immunogenic Hepatitis A Virus (HAV) peptide(s) - used to induce an
PT	Immune response to HAV in a mammal or to detect the presence of

PT antibodies against HAV in a mammal  
XX  
PS Claim 18; Page 112; 140pp; English.  
XX  
CC Peptides AAW42922-30 are immunogenic peptides corresponding to  
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are  
CC substantially similar to a portion of the amino acid sequence of the P2A  
CC protein of HAV corresponding to amino acids 792-980. The present peptide  
CC is derived from amino acids 922-941, and has a reactivity of 31.3% with  
CC acute sera. Compositions containing the peptides can be used to induce an  
CC immune response to HAV in a mammal. The peptides can also be used to  
CC detect the presence of antibodies against HAV in mammalian serum. The  
CC peptides can also be used to make an antibody against HAV by  
CC administering the peptide to a mammal.  
XX  
SQ Sequence 20 AA;  
Query Match 100.0%; Score 104; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.5e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KVNFPHGMDLDEETIANSKD 20  
1 KVNFPHGMDLDEETIANSKD 20  
Db  
RESULT 2  
AAB69444  
ID AAB69444 standard; Peptide; 21 AA.  
XX  
AC AAB69444;  
XX  
DT 20-APR-2001 (first entry)  
XX  
DE Synthetic HAV P2A peptide; SEQ ID NO: 44.  
XX  
KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;  
KW antigen; major structural capsid polypeptide; HAV antibody detection.  
XX  
OS Hepatitis A virus.  
OS Synthetic.  
XX  
PN WO200105824-A2.  
XX  
PD 25-JAN-2001.  
XX  
PF 14-JUL-2000; 2000MO-US19267.  
XX  
PS 15-JUL-1999; 99US-0144412.  
XX  
PI (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Fields HA, Khudyakov YE;  
XX  
DR WPI; 2001-112681/12.  
XX  
PT Synthetic peptides used as antigen sources for enzyme immunoassays  
XX detecting anti-hepatitis A virus and as vaccines -  
XX  
PS Claim 13; Page 96; 130pp; English.  
XX  
CC The present sequence is one of a number of synthetic peptides which are  
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides  
CC comprise antigenic epitopes of the major structural capsid polypeptides  
CC or non-structural polypeptides of HAV with one or more glutamine  
CC molecules at the carboxy end of the peptide. The peptides are used to  
CC detect the presence of antibodies against HAV in mammalian serum, to  
CC detect the presence of HAV in a human or animal through the binding of  
CC the peptide to an antibody, to detect acute phase infection by detecting  
CC IgM antibodies in mammalian serum and detecting convalescence in a  
CC mammal. The peptides are used to detect or quantify HAV antibodies in  
CC samples in clinical or research-based assays using immunoblotting,  
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,

CC tracking of radioactive or bioluminescent markers, chromatography or  
CC electrophoresis. The peptides are used to induce an immune response to  
CC HAV when administered to a human or animal. Glutamine at the carboxy  
CC end of the peptides enhances the IgM antibody reactivity.  
XX  
SQ Sequence 21 AA;  
Query Match 100.0%; Score 104; DB 22; Length 21;  
Best Local Similarity 100.0%; Pred. No. 1.6e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KVNFPHGMDLDEETIANSKD 20  
1 KVNFPHGMDLDEETIANSKD 20  
Db  
RESULT 3  
AAP60066  
ID AAP60066 standard; Protein; 2227 AA.  
XX  
AC AAP60066;  
XX  
DT 25-MAR-2003 (updated)  
DT 26-JUN-1991 (first entry)  
XX  
DE Sequence of viral L434 polypeptide encoded by the complete  
DE nucleotide sequence of the HAV genome.  
XX  
KW Diagnosis; vaccine; passive immunotherapy.  
XX  
OS Hepatitis A virus.  
XX  
FH Key  
FH Location/Qualifiers  
FT 1..245  
FT /label= P1.1A  
FT Region  
FT 246..491  
FT /label= 1B  
FT 492..836  
FT /label= 1C  
FT 837..980  
FT /label= P2.2A  
FT Region  
FT 981..1076  
FT /label= 2B  
FT 1077..1422  
FT /label= 2C  
FT 1423..1484  
FT /label= P3.3A  
FT Region  
FT 1485..1507  
FT /label= 3B  
FT 1508..1678  
FT /label= 3C  
FT 1679..2227  
FT /label= 3D  
XX  
XX EPI99480-A.  
XX  
XX 29-OCT-1986.  
XX  
XX 03-APR-1986; 86EP-0302465.  
XX  
XX 03-APR-1985; 85US-0719329.  
XX  
XX (CHIR ) CHIRON CORP.  
XX  
PI Dina D, Potter SJ, Vannest GA, Caput D;  
XX  
XX WPI; 1986-286213/44.  
XX  
XX N-PSDB; AAN60080.  
XX  
PT Hepatitis A virus nucleotide sequence and polypeptide - and use  
PT in prodn. of vaccines and diagnostic probes  
XX  
PS Claim 5; Fig 1; 18pp; English.



Query Match	Best Local Similarity	Score 104;	DB 7;	Length 2227;
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 KVNFPHGMLDEETIAANSKD 20			
DB	922 KVNFPHGMLDEETIAANSKD 941			
RESULT 4				
AA05697	AA05697 standard; protein; 2227 AA.			
XX	AA05697;			
XX	25-MAR-2003 (updated)			
DT	15-AUG-1990 (first entry)			
XX	Attenuated hepatitis A virus.			
XX	Hepatitis A virus; vaccine; attenuated.			
XX	Hepatitis A virus, strain HM-175.			
XX	Key			
FT	Location/Qualifiers			
FT	/label-VP4 = 1A			
FT	24..245			
FT	/label-VP2 = 1B			
FT	246..491			
FT	/label-VP3 = 1C			
FT	492..791			
FT	/label-VP1 = 1D			
FT	792..980			
FT	/label-2A			
FT	981..1087			
FT	/label-2B			
FT	1088..1422			
FT	/label-2C			
FT	1423..1496			
FT	/label-3A			
FT	1497..1519			
FT	/label-3B = VPg			
FT	1520..1738			
FT	/label-3C			
FT	1739..2227			
FT	/label-3D			
XX	Region			
XX	US4894228-A.			
XX	16-JAN-1990.			
XX	12-JUL-1988;	88US-0217824.		
XX	12-JUL-1988;	88US-0217824.		
XX	12-JUL-1988;	88US-0652967.		
XX	(USSH ) US DEPT HEALTH & HUMAN SERVICE.			
XX	Purcell RH, Titchhurst JR, Cohen JL, Emerson SU, Feinstein SM;			
XX	Daemer KU, Gust ID;			

[illegible]

XX 19-APR-1996; 960S-0015642.  
XX (USSH ) US SEC DEPT HEALTH.  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Emerson SU, Purcell RH, Raychaudhuri G;  
XX WPI; 1997-535850/49.  
DR N-PSDB; AAT93023.  
XX  
PT Human attenuated HAV genome containing simian HAV 2C gene - useful  
PT as vaccines against HAV infection  
XX  
PS Disclosure; Fig 13A-D; 66pp; English.  
XX  
CC This protein sequence is encoded by the human hepatitis A virus  
CC (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain  
CC HAV/7 is obtained by passage of HM-175 in African Green Monkey  
CC kidney cells. A claimed DNA construct (1) comprises a genome of  
CC HAV, where the genome is a human attenuated HAV genome in which a  
CC region of the 2C gene has been replaced by a corresponding region  
CC from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The  
CC region of the 2C gene from AGM-27 contained in the construct  
CC preferably encodes amino acids 120-328 of the 2C protein, amino  
CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA  
CC transcript of (1); (2) a cell transfected with (1) or the RNA  
CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the  
CC HAV of (3); and (5) a host cell containing the HAV of (3). (1) or  
CC its RNA transcript can be used as a vaccine for preventing HAV in  
CC a mammal. (1) or the RNA transcript can also be used to stimulate  
CC the production of protective antibodies in the mammal.  
XX  
SQ Sequence 2227 AA;  
OY Query Match 100.0%; Score 104; DB 18; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 KVNPHGMDLDEETANSKD 20  
922 KVNPHGMDLDEETANSKD 941  
OY RESULT 6  
AAB18607  
ID AAB18607 standard; Protein; 2227 AA.  
XX  
XX AAB18607;  
XX  
XX 15-JAN-2001 (first entry)  
XX  
DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.  
XX  
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.  
XX  
OS Hepatitis A virus.  
XX  
XX US6113912-A.  
XX  
XX PD 05-SEP-2000.  
XX  
XX PF 07-JUN-1995; 950S-0475886.  
XX  
XX PR 18-SEP-1992; 920S-0947338.  
XX  
XX PR 17-SEP-1993; 93MO-US08610.  
XX  
XX PR 10-MAR-1995; 950S-0397232.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;  
XX  
XX WPI; 2000-586464/55.  
XX  
DR

DR N-PSDB; AAA75476.  
XX  
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell  
XX line useful as vaccine for protecting humans against hepatitis A virus  
XX infection, has modified genome compared to wild type  
XX  
XX  
PS Disclosure; Fig 6A-K; 72pp; English.  
XX  
XX The present sequence is derived from a wild type hepatitis A virus  
XX (HAV) strain HM-174. The sequence is modified to produce HAV which  
XX are adapted to growth in the human fibroblast-like cell line MRC-5.  
XX The HAV is able to propagate in MRC-5 cells and retain appropriate  
XX attenuation. It is useful as a live vaccine for prophylaxis of  
XX hepatitis A in humans and other primates.  
XX  
SQ Sequence 2227 AA;  
OY Query Match 100.0%; Score 104; DB 21; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 KVNPHGMDLDEETANSKD 20  
922 KVNPHGMDLDEETANSKD 941  
OY RESULT 7  
AAB18608  
ID AAB18608 standard; Protein; 2227 AA.  
XX  
XX AAB18608;  
XX  
XX 15-JAN-2001 (first entry)  
XX  
XX Amino acid sequence of passage 35 Hepatitis A virus called P-35.  
XX  
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;  
XX P-35 virus.  
XX  
XX Hepatitis A virus.  
XX  
XX US6113912-A.  
XX  
XX PD 05-SEP-2000.  
XX  
XX PF 07-JUN-1995; 950S-0475886.  
XX  
XX PR 18-SEP-1992; 920S-0947338.  
XX  
XX PR 17-SEP-1993; 93MO-US08610.  
XX  
XX PR 10-MAR-1995; 950S-0397232.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;  
XX  
XX WPI; 2000-586464/55.  
XX  
XX N-PSDB; AAA75477.  
XX  
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell  
XX line useful as vaccine for protecting humans against hepatitis A virus  
XX infection, has modified genome compared to wild type  
XX  
XX  
PS Disclosure; Columns 67-78; 72pp; English.  
XX  
XX The present sequence is derived from passage 35 of a wild type  
XX hepatitis A virus (HAV) strain HM-174. The resulting virus is  
XX designated P-35 virus. The sequence is modified to produce HAV which  
XX are adapted to growth in the human fibroblast-like cell line MRC-5.  
XX The HAV is able to propagate in MRC-5 cells and retain appropriate  
XX attenuation. It is useful as a live vaccine for prophylaxis of  
XX hepatitis A in humans and other primates.  
XX  
SQ Sequence 2227 AA;  
OY

Query Match 100.0%; Score 104; DB 21; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNFPFGMDLEBIANSKD 20  
 DB 922 KVNFPFGMDLEBIANSKD 941

RESULT 8  
 AAB18609  
 ID AAB18609 standard; Protein; 2227 AA.  
 AC AAB18609;  
 XX  
 XX  
 DT 15-JAN-2001 (first entry)  
 XX  
 XX Amino acid sequence of live attenuated Hepatitis A virus 4380.  
 KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;  
 XX HAV 4380.  
 XX Hepatitis A virus.  
 OS  
 XX  
 XX US6113912-A.  
 PN  
 XX  
 PD 05-SEP-2000.  
 XX  
 XX 07-JUN-1995; 95US-0475886.  
 PF  
 XX  
 XX 18-SEP-1992; 92US-0947338.  
 PR  
 PR 17-SEP-1993; 93WO-US08610.  
 PR 10-MAR-1995; 95US-0397232.  
 XX  
 XX (USSR ) US DEPT HEALTH & HUMAN SERVICES.  
 PA  
 XX  
 XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;  
 PI  
 XX  
 XX WPI: 2000-586464/55.  
 DR  
 DR N-PSDB; AAA75478.  
 XX  
 XX Novel live hepatitis A virus adapted to growth in human fibroblast cell  
 PT line useful as vaccine for protecting humans against hepatitis A virus  
 PT infection, has modified genome compared to wild type  
 XX  
 XX  
 PS Disclosure: Columns 93-104; 72pp; English.

The present sequence is derived from a live attenuated hepatitis A virus (HAV) of the invention, designated HAV 4380. The sequence is produced by modifying wild type HAV strain HM-174. The HAV of the invention are adapted to growth in the human fibroblast-like cell line MRC-5. The HAV is able to propagate in MRC-5 cells and retain appropriate attenuation. It is useful as a live vaccine for prophylaxis of hepatitis A in humans and other primates.

XX  
 SQ Sequence 2227 AA;

Query Match 100.0%; Score 104; DB 21; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNFPFGMDLEBIANSKD 20  
 DB 922 KVNFPFGMDLEBIANSKD 941

RESULT 9  
 ABG31727  
 ID ABG31727 standard; Protein; 2227 AA.  
 AC ABG31727;  
 XX  
 XX

DT 29-NOV-2002 (first entry)  
 XX  
 XX Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.  
 DE  
 XX  
 XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.  
 KW  
 XX  
 XX Hepatitis A virus strain HM-175.  
 OS  
 XX  
 XX US6423318-B1.  
 PN  
 XX  
 PD 23-JUL-2002.  
 XX  
 XX 31-AUG-2000; 2000US-0653499.  
 PF  
 XX  
 XX 07-JUN-1995; 95US-0475886.  
 PR  
 PR 17-SEP-1993; 93US-0397232.  
 PR 17-SEP-1993; 93WO-US08610.  
 XX  
 XX (USSR ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (SMIR ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 XX Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;  
 PI  
 XX  
 XX WPI: 2002-680946/73.  
 DR  
 DR N-PSDB; ABS52787.  
 XX  
 XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth  
 PT in MRC-5 cells, useful for preparing a vaccine against HAV infection  
 XX  
 XX  
 PS Disclosure: Fig 6; 71pp; English.

The invention relates to a polynucleotide which encodes a hepatitis A virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell line). The polynucleotide is useful for preparing a vaccine against hepatitis A virus infection. This sequence represents a hepatitis A virus strain HM-175 polypeptide.

XX  
 SQ Sequence 2227 AA;

Query Match 100.0%; Score 104; DB 23; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNFPFGMDLEBIANSKD 20  
 DB 922 KVNFPFGMDLEBIANSKD 941

RESULT 10  
 ABG31728  
 ID ABG31728 standard; Protein; 2227 AA.  
 XX  
 XX ABG31728;  
 AC  
 XX  
 XX 29-NOV-2002 (first entry)  
 DT  
 XX  
 XX Hepatitis A virus mutant strain HM-175/7 (PHAV/7) polypeptide.  
 DE  
 XX  
 XX Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;  
 KW virucide; mutant; PHAV/7; mutlein.  
 XX  
 XX  
 XX Hepatitis A virus strain HM-175.  
 OS  
 XX  
 XX Synthetic.  
 FH  
 FH Key Location/Qualifiers  
 FT Misc-difference 963  
 FT /label= Wild-type Lys substituted by Arg  
 FT Misc-difference 764  
 FT /note= "Wild-type Glu substituted by Val"  
 FT Misc-difference 821  
 FT /note= "Wild-type Asn substituted by Ser"  
 FT Misc-difference 1052  
 FT /note= "Wild-type Ala substituted by Val"

```

FT      Misc-difference /note= "Wild-type Gly substituted by Ala"
FT      Misc-difference 1118 /note= "Wild-type Lys substituted by Met"
FT      Misc-difference 1151 /note= "Wild-type Glu substituted by Lys"
FT      Misc-difference 1163 /note= "Wild-type Phe substituted by Ser"
FT      Misc-difference 1277 /note= "Wild-type Val substituted by Ile"
FT      Misc-difference 1500 /note= "Wild-type His substituted by Tyr"
FT      Misc-difference 1805 /note= "Wild-type Asp substituted by Asn"
FT      Misc-difference 1930 /note= "Wild-type Ser substituted by Thr"
FT      Misc-difference /note= "Wild-type Ser substituted by Thr"
XX
XX      US6423318-B1.
XX
XX      23-JUL-2002.
XX
XX      31-AUG-2000; 2000US-0653499.
XX
PR      07-JUN-1995; 95US-0475886.
PR      17-SEP-1993; 93US-0397232.
PR      17-SEP-1993; 93WO-US08610.
XX
XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX      Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
XX      WPI; 2002-680946/73.
XX      N-PSDB; ABS52788.
XX
XX      New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX      in MRC-5 cells, useful for preparing a vaccine against HAV infection -
XX      Example 3; Column 67-76; 71pp; English.
XX
XX      The invention relates to a polynucleotide which encodes a hepatitis A
XX      virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX      line). The polynucleotide is useful for preparing a vaccine against
XX      hepatitis A virus infection. This sequence represents a hepatitis A virus
XX      mutant strain HM-175/7 (PHAV/7) polypeptide.
XX
XX      Sequence 2227 AA:
XX
SQ
Query Match          100.0%; Score 104; DB 23; Length 2227;
Fast Local Similarity 100.0%; Pred. NO. 2.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY      1 KVNFPHGMLDEETAAASKD 20
Qb      |||||||
        |||||||
        922 KVNFPHGMLDEETAAASKD 941
RESULTS 11
ABG31729
ID      ABG31729 standard; Protein: 2227 AA.
XX
XX      ABG31729;
XX
XX      29-NOV-2002 (first entry)
XX
XX      Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
XX
XX      Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
XX      HAV 4380.
XX
XX      Hepatitis A virus strain HM-175.
XX
XX      US6423318-B1

```

```

XX      23-JUL-2002.
PD
XX      31-AUG-2000; 2000US-0653499.
PE
XX      07-JUN-1995;   95US-0475886.
PR      17-SEP-1993;   93US-0397232.
PR      17-SEP-1993;   93WO-US08610.
XX
XX      (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PA      (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX      Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;
DR
XX      WPI: 2002-680946/73.
DR      N-PSDB; AAB552789.
XX
XX      New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
PT      in MRC-5 cells, useful for preparing a vaccine against HAV infection -
PT
XX      Disclosure; Column 93-104; 71pp; English.
PS
XX
XX      The invention relates to a polynucleotide which encodes a hepatitis A
CC      virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
CC      line). The polynucleotide is useful for preparing a vaccine against
CC      hepatitis A virus infection. This sequence represents an attenuated
CC      hepatitis A virus 4830 polypeptide.
XX
XX      Sequence    2227 AA;
SQ
XX
Query Match          100.0%; Score 104; DB 23; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.8e-08;
Matches    20; Conservative    0; Mismatches    0; Indels    0; Gaps    0;
QY      1 KVNPHGMLDDEETRANSKD 20
        |||||||
Db       922 KVNPHGMLDDEETRANSKD 941
XX
RESULT 12
AAE19899
ID      AAE19899 standard; Protein: 2227 AA.
AC
XX      AAE19899;
XX
XX      18-JUN-2002 (first entry)
DT
XX
XX      Hepatitis A virus (HAV) protein.
DE
XX
XX      Hepatitis A virus; HAV; Infection; virucide; fungicide; antibacterial;
KW      cyostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
XX      Hepatitis A virus.
OS
XX      WO200213855-A2.
PN
XX      21-FEB-2002.
XX
XX      15-AUG-2001; 2001WO-IB01808.
PF
XX
XX      17-AUG-2000; 2000US-225767P.
PR      29-AUG-2000; 2000US-229175P.
PR      03-NOV-2000; 2000US-0705547.
XX
XX      (TRIP-) TRIPEP AB.
PA
XX      Salberg M, Hultgren C;
PI
XX      WPI: 2002-241837/29.
DR      N-PSDB; AAD31766.
DR
XX
XX      Vaccine compositions for treating and preventing disease, preferably
PT      hepatitis C virus infection, comprises ribavirin and antigen that has
XX

```

PT epitope present in hepatitis C virus -  
XX  
PS Claim 11; Page 82-87; 120pp; English.  
XX  
CC The invention relates to a composition comprising ribavirin and an  
antigen preferably non structural 3 protein (NS3)/4A fragment of  
hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV  
sequence. The composition is useful for enhancing an immune response to  
a hepatitis C antigen in humans, domestic, sport or pet species and as  
vaccines for treating and preventing HCV infections. The composition is  
also useful for treating viral, bacterial, fungal diseases and cancer.  
CC The present sequence is hepatitis A virus (HAV) protein.  
XX  
SQ Sequence 2227 AA;  
Query Match 100.0%; Score 104; DB 23; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KVNPFHGMIDLEIANSKD 20  
|||||  
922 KVNPFHGMIDLEIANSKD 941  
DB  
RESULT 13  
ABU08639  
ID ABU08639 standard; Protein; 2227 AA.  
XX  
AC ABU08639;  
XX  
DT 03-JUN-2003 (first entry)  
XX  
DE Wild type human hepatitis A virus strain HM-175.  
XX  
KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;  
vaccine; MRC-5 cell; hepatitis infection.  
XX  
OS Hepatitis A virus strain HM-175.  
XX  
PN US2002176869-A1.  
XX  
PD 28-NOV-2002.  
XX  
PF 29-APR-2002; 2002US-0135988.  
XX  
PR 07-JUN-1995; 95US-0475886.  
PR 31-AUG-2000; 2000US-0653499.  
PR 18-SEP-1992; 92US-0947338.  
PR 17-SEP-1993; 93WO-0508610.  
PR 17-APR-1995; 95US-0397232.  
XX  
PA (FUNK/) FUNKHOUSER A W.  
PA (EMER/) EMERSON S U.  
PA (PURC/) PURCELL R H.  
PA (DHON/) D'HONDT E.  
XX  
PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;  
XX  
DR MPI. 2003-352605/33.  
DR N-PSDB; ABX93473.  
XX  
PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,  
PT useful in vaccines for protecting primates against hepatitis infection  
and disease -  
XX  
PS Disclosure; Fig 6; 70pp; English.  
XX  
CC The invention describes a live hepatitis A virus (HAV) adapted to growth  
in MRC-5 cells. The HAV and compositions comprising the HAV are useful as  
a vaccine for protecting primates against hepatitis infection and  
disease. This is the amino acid sequence of wild type human hepatitis A  
virus strain HM-175.  
CC  
XX

SQ Sequence 2227 AA;  
Query Match 100.0%; Score 104; DB 24; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KVNPFHGMIDLEIANSKD 20  
|||||  
922 KVNPFHGMIDLEIANSKD 941  
DB  
RESULT 14  
ABU08640  
ID ABU08640 standard; Protein; 2227 AA.  
XX  
AC ABU08640;  
XX  
DT 03-JUN-2003 (first entry)  
XX  
DE Attenuated (pass35) hepatitis A virus strain HM-175.  
XX  
KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;  
vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.  
XX  
OS Hepatitis A virus strain HM-175.  
XX  
PN US2002176869-A1.  
XX  
PD 28-NOV-2002.  
XX  
PF 29-APR-2002; 2002US-0135988.  
XX  
PR 07-JUN-1995; 95US-0475886.  
PR 31-AUG-2000; 2000US-0653499.  
PR 18-SEP-1992; 92US-0947338.  
PR 17-SEP-1993; 93WO-0508610.  
PR 17-APR-1995; 95US-0397232.  
XX  
PA (FUNK/) FUNKHOUSER A W.  
PA (EMER/) EMERSON S U.  
PA (PURC/) PURCELL R H.  
PA (DHON/) D'HONDT E.  
XX  
PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;  
XX  
DR MPI. 2003-352605/33.  
DR N-PSDB; ABX93474.  
XX  
PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,  
PT useful in vaccines for protecting primates against hepatitis infection  
and disease -  
XX  
PS Example 3; Fig 6; 70pp; English.  
XX  
CC The invention describes a live hepatitis A virus (HAV) adapted to growth  
in MRC-5 cells. The HAV and compositions comprising the HAV are useful as  
a vaccine for protecting primates against hepatitis infection and  
disease. This is the amino acid sequence of an attenuated (pass 35)  
human hepatitis A virus strain HM-175.  
CC  
XX  
SQ Sequence 2227 AA;  
Query Match 100.0%; Score 104; DB 24; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KVNPFHGMIDLEIANSKD 20  
|||||  
922 KVNPFHGMIDLEIANSKD 941  
DB  
RESULT 15  
ABU08641

ID AB008641 standard; Protein: 2227 AA.  
 AC AB008641;  
 XX  
 XX  
 DT 03-JUN-2003 (first entry)  
 XX  
 DE Attenuated hepatitis A virus (4380) strain HM-175.  
 XX  
 KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;  
 XX vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.  
 OS Hepatitis A virus strain HM-175.  
 PN US202176869-A1.  
 PD 28-NOV-2002.  
 PD 29-APR-2002; 2002US-0135988.  
 PF 07-JUN-1995; 95US-0475886.  
 PR 31-AUG-2000; 2000US-0653499.  
 PR 18-SEP-1992; 92US-0947338.  
 PR 17-SEP-1993; 93WO-US08610.  
 PR 17-APR-1995; 95US-0397232.  
 XX  
 PA (FUNK/) FUNKHOUSER A W.  
 PA (EMER/) EMERSON S U.  
 PA (PURC/) PURCELL R H.  
 PA (DHON/) D'HONDT E.  
 XX  
 PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;  
 DR WPI: 2003-352605/33.  
 DR N-PSDB: ABX93475.  
 XX  
 PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,  
 PT useful in vaccines for protecting primates against hepatitis infection  
 PT and disease -  
 XX  
 PS Disclosure; Page 45-51; 70pp; English.  
 XX  
 CC The invention describes a live hepatitis A virus (HAV) adapted to growth  
 CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as  
 CC a vaccine for protecting primates against hepatitis infection and  
 CC disease. This is the amino acid sequence of an attenuated human  
 CC hepatitis A virus (4380) strain HM-175.  
 XX  
 SO Sequence 2227 AA;  
 XX  
 XX Query Match 100.0%; Score 104; DB 24; Length 2227;  
 XX Best Local Similarity 100.0%; Pred. No. 2.8e-08;  
 XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 KVNPPHGMIDLEETANSKD 20  
 DB 922 KVNPPHGMIDLEETANSKD 941

Search completed: October 1, 2003, 09:56:43  
 Job time : 44.5294 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 12.8235 Seconds  
(without alignments)  
149.988 Million cell updates/sec

Title: US-09-171-432a-44

Perfect score: 104  
Sequence: 1 KVNPPHGMJDLIEIANSKD 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	2227	1 GNNYHM	genome polyprotein
2	104	100.0	2227	1 GNNYHR	genome polyprotein
3	104	100.0	2227	1 GNNYMK	genome polyprotein
4	104	100.0	2227	1 GNNYHB	genome polyprotein
5	101	97.1	2230	1 GNNYSA	genome polyprotein
6	47	45.2	333	2 A12485	hypothetical prote
7	47	45.2	387	2 B86669	RecA protein [Impo
8	47	45.2	506	1 S58522	glycine-tRNA 119s
9	47	45.2	1060	2 F88710	protein C01G5.4 [1
10	47	45.2	1079	2 T30996	hypothetical prote
11	46	44.2	259	2 F82087	deoxyribose-phosph
12	46	44.2	444	2 C89768	conserved hypotet
13	46	44.2	677	2 H64574	DNA topoisomerase
14	45	43.3	195	2 T08812	probable succinate
15	45	43.3	383	2 F96582	hypothetical prote
16	45	43.3	417	2 A44529	succinate-CoA liiga
17	45	43.3	852	2 B72685	hypothetical prote
18	45	43.3	935	2 T19011	hypothetical prote
19	45	43.3	1213	2 T19835	hypothetical prote
20	44	42.3	223	2 T37962	hypothetical prote
21	44	42.3	344	2 D97761	hypothetical prote
22	44	42.3	500	2 S50508	ANP1 protein - yea
23	43.5	41.8	373	2 A69773	hypothetical prote
24	43	41.3	140	2 E81659	conserved hypotet
25	43	41.3	141	2 G71501	hypothetical prote
26	43	41.3	257	2 AF2592	2-deoxyribose-5-ph
27	43	41.3	309	2 H97374	2-deoxyribose-5-ph
28	43	41.3	359	2 H71089	hypothetical prote
29	43	41.3	327	2 T00876	probable glycerate

30	43	41.3	440	2 B71858	adenylosuccinate 1
31	43	41.3	440	2 H64658	adenylosuccinate 1
32	43	41.3	609	2 AB0955	glutamine-fructose
33	43	41.3	917	2 T21870	hypothetical prote
34	42.5	40.9	599	2 A86810	1-deoxyxylulose-5-
35	42	40.4	222	2 C82243	conserved hypotet
36	42	40.4	229	2 S77449	hydrogenase expres
37	42	40.4	259	2 D91296	2-deoxyribose-5-ph
38	42	40.4	290	2 S19426	hypothetical prote
39	42	40.4	309	2 T33259	hypothetical prote
40	42	40.4	342	2 C72313	hypothetical prote
41	42	40.4	365	2 B48945	recombination prote
42	42	40.4	384	2 E82088	conserved hypotet
43	42	40.4	495	2 D64578	conserved hypotet
44	42	40.4	539	2 G70520	probable csp prote
45	42	40.4	539	2 G70520	probable csp prote

#### ALIGNMENTS

##### RESULT 1

GNNYHM genome polyprotein - human hepatitis A virus (strain HM-175, wild type)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core  
B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 16-Jul-1999

C:Accession: A25981

R:Chen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with

A:Reference number: A25981, NCID:87061253, PMID:3023706

A:Accession: A25981

A:Molecule type: genomic RNA

A:Residues: 1-2227 <COR>

A:Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583

C:Superfamily: hepatitis A virus genome polyprotein

C:Keywords: coat protein; core protein; cysteine proteinase; hydroxylase; nucleotidyltr

F:1-23/Product: coat protein 1A #status predicted <VP1>

F:24-245/Product: coat protein 1B #status predicted <VP2>

F:446-491/Product: coat protein 1C #status predicted <VP3>

F:492-791/Product: coat protein 1D #status predicted <VP4>

F:792-980/Product: core protein 2A #status predicted <C2A>

F:981-1087/Product: core protein 2B #status predicted <C2B>

F:1088-1422/Product: core protein 2C #status predicted <C2C>

F:1423-1496/Product: protein 3A #status predicted <C3A>

F:1497-1519/Product: protein 3B #status predicted <C3B>

F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.5e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPPHGMJDLIEIANSKD 20  
Db 922 KVNPPHGMJDLIEIANSKD 941

##### RESULT 2

GNNYHR genome polyprotein - human hepatitis A virus

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core

NA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Jul-1999

C:Accession: A03903

R:Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N

Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985

A:Title: Primary structure and gene organization of human hepatitis A virus.

A:Reference number: A03903; MUID:85190549; PMID:2986127  
A:Accession: A03903  
A:Molecule type: genomic RNA  
A:Residues: 1-2227 <NAU>  
A:Cross-references: GB:R02990; NID:g3329596; PIDN:AAA45471.1; PID:g3329597  
C:Superfamily: hepatitis A virus genome polypeptide  
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans  
F:1-245/Product: coat protein 1A #status predicted <C1A>  
F:246-491/Product: coat protein 1B #status predicted <C1B>  
F:492-836/Product: coat protein 1C #status predicted <C1C>  
F:837-980/Product: core protein 2A #status predicted <C2A>  
F:981-1076/Product: core protein 2B #status predicted <C2B>  
F:1077-1422/Product: core protein 2C #status predicted <C2C>  
F:1423-1484/Product: protein 3A #status predicted <C3A>  
F:1485-1507/Product: protein 3B #status predicted <C3B>  
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>  
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMLDEETIANSKD 20  
|||||  
Db 922 KVNPHGMLDEETIANSKD 941

RESULT 3  
GNNYMB  
genome polypeptide - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core pro  
NA polymerase (EC 2.7.7.48), protein 3D  
C:Species: human hepatitis A virus  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1988 #sequence revision 30-Jun-1988 #text\_change 16-Jul-1999  
C:Accession: A94149; A25914; A94508  
R:Cohen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstein, S.M.; Purcell, R  
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987  
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison with  
A:Reference number: A94149; MUID:87175701; PMID:3031686  
A:Accession: A94149  
A:Status: nucleic acid sequence not shown  
A:Molecule type: genomic RNA  
A:Residues: 1-2227 <COH>  
A:Cross-references: EMBL:M1632; NID:g3329594; PIDN:AAA45471.1; PID:g3329595  
A:Note: Submitted to GenBank, August 1987  
C:Superfamily: hepatitis A virus genome polypeptide  
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltrans  
F:1-245/Product: coat protein 1A #status predicted <P1A>  
F:246-491/Product: coat protein 1B #status predicted <P1B>  
F:492-836/Product: coat protein 1C #status predicted <P1C>  
F:837-980/Product: core protein 2A #status predicted <P2A>  
F:981-1076/Product: core protein 2B #status predicted <P2B>  
F:1077-1422/Product: core protein 2C #status predicted <P2C>  
F:1423-1484/Product: protein 3A #status predicted <P3A>  
F:1485-1507/Product: protein 3B #status predicted <P3B>  
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>  
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMLDEETIANSKD 20  
|||||  
Db 922 KVNPHGMLDEETIANSKD 941

RESULT 4  
GNNYMB  
genome polypeptide - human hepatitis A virus (strain HMB)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro  
Vpg; protein 3A; RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D

C:Species: human hepatitis A virus  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1991 #sequence revision 30-Jun-1991 #text\_change 10-May-1996  
C:Accession: J50303  
R:Paul, A.V.; Tada, H.; von der Helm, K.; Wessel, T.; Klein, R.; Wimmer, E.; Delnath  
virus Res. 8, 153-171, 1987  
A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (iso  
A:Reference number: J50303; MUID:88045071; PMID:2823500  
A:Accession: J50303  
A:Molecule type: genomic RNA  
A:Residues: 1-2227 <PAU>  
A:Cross-references: EMBL:M20273  
C:Superfamily: hepatitis A virus genome polypeptide  
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; h  
F:1-23/Product: coat protein 1A #status predicted <VP1>  
F:24-246/Product: coat protein 1B #status predicted <VP2>  
F:247-491/Product: coat protein 1C #status predicted <VP3>  
F:492-836/Product: coat protein 1D #status predicted <VP4>  
F:837-980/Product: core protein 2A #status predicted <P2A>  
F:981-1108/Product: core protein 2B #status predicted <P2B>  
F:1109-1438/Product: core protein 2C #status predicted <P2C>  
F:1439-1496/Product: protein 3A #status predicted <P3A>  
F:1497-1519/Product: genome-linked protein vpg #status predicted <VP6>  
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>  
F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 100.0%; Score 104; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.5e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMLDEETIANSKD 20  
|||||  
Db 922 KVNPHGMLDEETIANSKD 941

RESULT 5  
GNNYSA  
genome polypeptide - simian hepatitis A virus (strain AGM-27)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core  
C:Species: simian hepatitis A virus  
C:Date: 31-Mar-1992 #sequence revision 31-Mar-1992 #text\_change 16-Jun-2000  
C:Accession: A30470; S04885; S03965  
R:Tsarev, S.A.  
submitted to JFID, April 1991  
A:Reference number: A30470  
A:Accession: A30470  
A:Molecule type: genomic RNA  
A:Residues: 1-2230 <TSA>  
A:Cross-references: GB:D00924; NID:g222597; PIDN:BA00766.1; PID:g222598  
R:Tsarev, S.A.; Emerson, S.O.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.  
J. Gen. Virol. 72, 1677-1683, 1991  
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure  
A:Reference number: J01080; MUID:91311420; PMID:1649901  
A:Contents: annotation  
A:Note: neither amino acid nor nucleotide sequence is given  
R:Balayan, M.S.; Kusov, Y.Y.; Andjapardide, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh  
submitted to the EMBL Data Library, May 1989  
A:Reference number: S04885  
A:Accession: S04885  
A:Molecule type: genomic RNA  
A:Residues: 1750-2164 <BAL1>  
A:Cross-references: EMBL:X15461; NID:961971; PIDN:CA03490.1; PID:g930268  
R:Balayan, M.S.; Kusov, Y.Y.; Andjapardide, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh  
FEBS Lett. 247, 425-428, 1989  
A:Title: Variations in genome fragments coding for RNA polymerase in human and simian  
A:Reference number: S03965; MUID:89232166; PMID:2341023  
A:Accession: S03965  
A:Molecule type: genomic RNA  
A:Residues: 1960-2164 <BAL2>  
A:Cross-references: EMBL:X15461  
C:Superfamily: hepatitis A virus genome polypeptide  
C:Keywords: coat protein; core protein; polypeptide  
F:1-27/Product: coat protein 1A #status predicted <C1A>



F:28-249/Product: coat protein 1B #status predicted <C1B>  
 F:350-495/Product: coat protein 1C #status predicted <C1C>  
 F:496-795/Product: coat protein 1D #status predicted <C1D>  
 F:796-984/Product: core protein 2A #status predicted <C2A>  
 F:985-1091/Product: core protein 2B #status predicted <C2B>  
 F:1092-1426/Product: core protein 2C #status predicted <C2C>  
 F:1427-1498/Product: protein 3A #status predicted <P3A>  
 F:1499-1521/Product: protein 3B #status predicted <P3B>  
 F:1522-1741/Product: protein 3C #status predicted <P3C>  
 F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 97.1%; Score 101; DB 1; Length 2230;  
 Best Local Similarity 95.0%; Pred. No. 4, 6e-08;  
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMLDEEIANSKD 20  
 DB 926 KVNPHGMLDEEIANSKD 945

RESULT 6

hypothetical protein all7065 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120a1  
 C:Species: Nostoc sp. PCC 7120  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
 C:Accession: A12485  
 R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kultz, T.; Sasamoto, S.; Matanabe, A.; Iriguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
 A:Reference number: AB1807; MID:21595285; PMID:11759840  
 A:Status: preliminary  
 A:Accession: A12485  
 A:Molecule type: DNA  
 A:Residues: 1-333 <KUR>  
 A:Cross-references: GB:BA000020; PIDN:BA078149.1; PID:g17135603; GSPDB:GN00180  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: all7065  
 A:Genome: plasmid

Query Match 45.2%; Score 47; DB 2; Length 333;  
 Best Local Similarity 40.0%; Pred. No. 5, 6;  
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

OY 1 KVNPHGMLDEEIANSKD 20  
 57 KSGYPRGKITRDEVANVOD 76

RESULT 7

Reca protein [imported] - Lactococcus lactis subsp. lactis (strain IL1403)  
 C:Species: Lactococcus lactis subsp. lactis  
 C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
 C:Accession: B86669  
 R:Boletijn, A.; Winkler, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weisenbach, J.; Ehrlich, S.  
 Genome Res. 11, 731-753, 2001  
 A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s.s.  
 A:Reference number: AB6625; MID:21235186; PMID:11337471  
 A:Accession: B86669  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-387 <STO>  
 A:Cross-references: GB:AE005176; PID:g12723223; PIDN:AAK04452.1; GSPDB:GN00146  
 A:Experimental source: strain IL1403  
 C:Genetics:  
 A:Gene: reca  
 C:Superfamily: recombination protein reca

Query Match 45.2%; Score 47; DB 2; Length 387;  
 Best Local Similarity 45.0%; Pred. No. 6, 7;

Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
 OY 1 KVNPHGMLDEEIANSKD 20  
 DB 340 KVRTANGLDEAEVATETD 359

RESULT 8

glycyl-tRNA ligase (EC 6.1.1.14) - Thermus aquaticus  
 S58522  
 N:Alternate names: glycyl-tRNA synthetase  
 C:Species: Thermus aquaticus  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Jun-2002  
 C:Accession: S58522  
 R:Logan, D.T.; Mazauric, M.H.; Kern, D.; Moras, D.  
 EMBO J. 14, 4156-4167, 1995  
 A:Title: Crystal structure of glycyl-tRNA synthetase from Thermus thermophilus.  
 A:Reference number: S58522; MID:96016187; PMID:7556056  
 A:Accession: S58522  
 A:Status: not compared with conceptual translation  
 A:Molecule type: DNA  
 A:Residues: 1-506 <LOG>  
 A:Experimental source: strain HB8  
 A:Note: the source is designated as Thermus thermophilus  
 C:Superfamily: Mycoplasma genitalium glycyl-tRNA ligase  
 C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis

Query Match 45.2%; Score 47; DB 1; Length 506;  
 Best Local Similarity 75.0%; Pred. No. 9, 2;  
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 FPHGMLDEEIA 15  
 DB 299 FPHGMLDEEIA 310

RESULT 9

protein C01G5.4 [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 C:Accession: F88710  
 R:anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio  
 A:Reference number: A75000; MID:99069613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/genC/elegans/ and www.sanger.ac.uk/Projects/C\_elegans/  
 A:Accession: F88710  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-1060 <STO>  
 A:Cross-references: GB:chr\_IV; PIDN:AA837736.1; PID:g1208856; GSPDB:GN00022; CESP:C01  
 C:Genetics:  
 A:Gene: C01G5.4  
 A:Map position: 4

Query Match 45.2%; Score 47; DB 2; Length 1060;  
 Best Local Similarity 36.8%; Pred. No. 22;  
 Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

OY 2 VNFPHGMLDEEIANSKD 20  
 DB 391 IGFSRGITLDKHYAGDARD 409

RESULT 10

hypothetical protein C01G5.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C>Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T30996  
 R:Bradehaw, H.; Stellies, L.

submitted to the EMBL Data Library, August 1999  
A:Description: The sequence of C. elegans cosmid C01G5.

A:Reference number: Z20956

A:Accession: T30996

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 11079 <BR>

A:Cross-references: EMBL:U50068; PIDN:AB37736.2

A:Experimental source: Strain Bristol N2

C:Genetics:

A:Map position: IV

A:Introns: 10/1; 31/1; 75/1; 108/1; 144/1; 927/2; 1056/2

A:Note: C01G5.4

Query Match 45.2%; Score 47; DB 2; Length 1079;  
Best Local Similarity 36.8%; Pred. No. 22;

Matches 7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNPPHGMIDLEIANSKD 20

410 IGFSGRIIDLNHVGADMD 428

RESULT 11  
F82087

deoxyribose-phosphate aldolase VC2350 [imported] - Vibrio cholerae (strain N16961 serogr  
C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: F82087

R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: F82087

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-259 <HE>

A:Cross-references: GB:AE004305; GB:AE003852; NID:99656912; PIDN:AA95493.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC2350

A:Map position: 1

C:Superfamily: deoxyribose-phosphate aldolase

Query Match 44.2%; Score 46; DB 2; Length 259;  
Best Local Similarity 52.9%; Pred. No. 6.2;

Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

3 NFPHGMIDLEIANSK 19

75 NFPHGMIDLEIANSK 91

Db

RESULT 12

C89768

conserved hypothetical protein SA0083 [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 22-Oct-2001

C:Accession: C89768

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogud

ma, A.; Mizutani, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;

C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.

Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: C89768

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-444 <KUR>

A:Cross-references: GB:BA000018; PID:q13700003; PIDN:BA041302.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:  
A:Gene: SA0083

Query Match 44.2%; Score 46; DB 2; Length 444;  
Best Local Similarity 52.6%; Pred. No. 12;

Matches 10; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 2 VNPPHGMIDLEIANSKD 20

384 VNPPHGMIDLEIANSKD 402

Db

RESULT 13

H64574

DNA topoisomerase I - Helicobacter pylori (strain 26695)

C:Species: Helicobacter pylori

C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 08-Oct-1999

C:Accession: H64574

R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.

Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKe

son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Wathey,

Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser,

A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.

A:Reference number: A64520; MUID:97394467; PMID:9252185

A:Accession: H64574

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-677 <TOM>

A:Cross-references: GB:AE000559; GB:AE000511; NID:q2313536; PIDN:AD07502.1; PID:q231

C:Superfamily: DNA topoisomerase I

Query Match 44.2%; Score 46; DB 2; Length 677;  
Best Local Similarity 31.6%; Pred. No. 19;

Matches 6; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 KVNPPHGMIDLEIANSK 19

346 RITHPHALKDEKVCSDAK 364

Db

RESULT 14

T08812

probable succinate-CoA ligase (GDP-forming) (EC 6.2.1.4) beta chain - human (fragment

N:Alternate names: protein DKFP586M2023.1; succinyl-CoA synthetase (GDP-forming) bet

C:Species: Homo sapiens (man)

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 03-Jun-2002

C:Accession: T08812

R:Angorge, W.; Winkler, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16472

A:Accession: T08812

A:Molecule type: mRNA

A:Residues: 1-195 <ANS>

A:Cross-references: EMBL:AL050226

A:Experimental source: adult uterus; clone DKFP586M2023

C:Genetics:

A:Note: DKFP586M2023.1

C:Function: catalyzes the formation of succinyl-CoA from succinate with concomitan

C:Superfamily: succinate-CoA ligase (ADP-forming) beta chain

C:Keywords: acid-thiol ligase; coenzyme A; mitochondrion

Query Match 43.3%; Score 45; DB 2; Length 195;  
Best Local Similarity 42.1%; Pred. No. 6.6;

Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNPPHGMIDLEIANSKD 20

65 VGSPOGVDIEVAASNP 83

Db

RESULT 15



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 : Search time 6.58824 Seconds

(without alignments)  
142.760 Million cell updates/sec

Title: US-09-171-432a-44

Perfect score: 104

Sequence: 1 KVNPPHGMJDEEIANSKD 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08

Maximum Match 1008  
Listing first 45 summaries

Database : SwissProt\_41:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	104	100.0	2226 1	P26581 hepatitis a
2	104	100.0	2226 1	P26582 hepatitis a
3	104	100.0	2227 1	P08617 hepatitis a
4	104	100.0	2227 1	P06441 hepatitis a
5	104	100.0	2227 1	P13901 hepatitis a
6	101	97.1	2230 1	P14553 simian hepa
7	100	96.2	2226 1	P26580 hepatitis a
8	100	95.2	387 1	P01840 lactococcus
9	47	45.2	404 1	Q92218 mus muscul
10	47	45.2	505 1	P56206 thermus the
11	46	44.2	259 1	Q9K217 vibrio chol
12	45	43.3	432 1	Q96199 homo sapien
13	45	43.3	433 1	P53590 sus scrofa
14	44	42.3	500 1	P32629 saccharomy
15	44	41.3	98 1	O75711 homo sapien
16	43	41.3	259 1	O8U109 agrobacteri
17	43	41.3	357 1	O57429 gallus gall
18	43	41.3	440 1	Q92K22 helicobacte
19	43	41.3	440 1	P56468 helicobacte
20	43	41.3	608 1	GIMS_SALTI
21	42	40.4	214 1	G1P1_POAST
22	42	40.4	259 1	P08236 pharbitis n
23	42	40.4	290 1	Q8X236 escherichia
24	41.5	39.9	406 1	Q8QD19 proteus sp.
25	41	39.4	193 1	P44339 haemophilus
26	41	39.4	218 1	O92966 rickettsia
27	41	39.4	400 1	O43884 anabaena az
28	41	39.4	400 1	P12623 anabaena sp
29	41	39.4	502 1	Q9MUG6 mesostigma
30	41	39.4	552 1	O8XMP3 clostridium
31	40.5	38.9	237 1	P17798 agrobacteri
32	40.5	38.9	434 1	O05973 rickettsia
33	40.5	38.9	1286 1	P20504 vaccinia vi

34	40.5	38.9	1286 1	RPOL_VAVR	P33053 variola vir
35	40.5	38.9	1287 1	RPOL_VACCV	P07392 vaccinia vi
36	40	38.5	126 1	YF81_XYLF	O9P822 xylella fas
37	40	38.5	310 1	YCCK_BACST	P46905 bacillus fas
38	40	38.5	353 1	UBP2_MOUSE	O88623 mus musculu
39	40	38.5	404 1	ISCS_NEIMA	O9JLX0 neisseria m
40	40	38.5	446 1	SYG_MYCE	P47493 mycoplasma
41	40	38.5	507 1	YIR4_YEAST	P40486 saccharomyc
42	40	38.5	558 1	PILE_NEIGO	P37094 saccharomyc
43	40	38.5	595 1	KMT2_HUMAN	O9H1R3 homo sapien
44	40	38.5	605 1	UBP2_HUMAN	O75604 homo sapien
45	40	38.5	646 1	KDBE_SCHPO	O10364 schizosacch

## ALIGNMENTS

RESULT 1  
POLG\_HPAV4 STANDARD: PRT: 2226 AA.  
AC P26581;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
DE P3D (EC 2.7.7.48).  
OS Hepatitis A virus (strain 43C).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12095;  
RN [1]  
RX MEDLINE=91162758; PubMed=1705995;  
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,  
RA "Antigenic and genetic variation in cytopathic hepatitis A virus  
RT variants arising during persistent infection: evidence for genetic  
RT recombination.";  
RT J. Virol. 65:2056-2065(1991).  
RL -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
[RNA](n).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
CC or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).  
CC -----  
CC EMBL: M59809; AAA5469.1; -  
CC MEROPS: C03.005; -  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR007094; RNA\_pol\_PSVIR.  
DR Pfam: PF00680; RNA\_dep\_RNA\_pol\_PSVIR.  
DR Pfam: PF00910; RNA\_helicase; 1.  
KW Polyprotein; Coat protein; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 794  
FT CHAIN 795 900  
FT CHAIN 901 1087  
FT CHAIN 1088 1422  
FT CORE PROTEIN P2A.  
FT CORE PROTEIN P2B.  
FT CORE PROTEIN P2C.

FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.  
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.  
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.  
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.  
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B09BF75 CRC64;  
Query Match 100.0%; Score 104; DB 1; Length 2226;  
Best Local Similarity 100.0%; Pred. No. 2.3e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KVNPHGMLDEETIANSKD 20  
Db 922 KVNPHGMLDEETIANSKD 941  
RESULT 2  
POLG\_HPAV8 STANDARD; PRT: 2226 AA.  
ID POLG\_HPAV8  
AC P26582;  
DE 01-AUG-1992 (Rel. 23, Created)  
DE 01-AUG-1992 (Rel. 23, Last sequence update)  
DE 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Genome polypotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)]  
DE Hepatitis A virus (strain 18f).  
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12096;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91162758; PubMed=1705995;  
RA Lemmon S.M., Murphy P.C., Shields P.A., Ping L.H., Felstone S.M.,  
RA Cromeans T., Jansen R.W.;  
RT "Antigenic and genetic variation in cytopathic hepatitis A virus  
RT variants arising during persistent infection: evidence for genetic  
RT recombination."  
RT J. Virol. 65:2056-2065(1991).  
RL J. Virol. 65:2056-2065(1991).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +  
CC (RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: M59808; AAA45467.1; -  
CC PDB: 1Q47; 15-MAY-00.  
DR MEROPS: C03.005.-  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR007094; RNA\_pol\_PSVir.  
DR Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam: PF00910; RNA\_helicase; 1.  
KW Polypeptin; Coat protein; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 794  
FT CHAIN 795 900  
FT CHAIN 901 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1495  
COAT PROTEIN VP4 (PIA).  
COAT PROTEIN VP2 (PIB).  
COAT PROTEIN VP3 (PIC).  
COAT PROTEIN VP1 (PID).  
CORE PROTEIN P2A.  
CORE PROTEIN P2B.  
CORE PROTEIN P2C.  
PROBABLE PROTEIN P3A.

FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.  
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.  
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.  
SQ SEQUENCE 2226 AA; 251292 MW; 249644A63396C8D6B CRC64;  
Query Match 100.0%; Score 104; DB 1; Length 2226;  
Best Local Similarity 100.0%; Pred. No. 2.3e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KVNPHGMLDEETIANSKD 20  
Db 922 KVNPHGMLDEETIANSKD 941  
RESULT 3  
POLG\_HPAV8 STANDARD; PRT: 2227 AA.  
ID POLG\_HPAV8  
AC P08617; P06443; Q81082;  
DE 01-AUG-1988 (Rel. 08, Created)  
DE 01-AUG-1988 (Rel. 08, Last sequence update)  
DE 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Genome polypotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)]  
DE Hepatitis A virus (strain HM-175).  
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12098;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Wild type;  
RX MEDLINE=87061253; PubMed=3023706;  
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,  
RA Baroudy B.M.;  
RT "Complete nucleotide sequence of wild-type hepatitis A virus:  
RT comparison with different strains of hepatitis A virus and other  
RT picornaviruses."  
RT J. Virol. 61:50-59(1987).  
RL [2]  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Attenuated;  
RX MEDLINE=87175701; PubMed=3031686;  
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Felstone S.M.,  
RA Purcell R.H.;  
RT "Complete nucleotide sequence of an attenuated hepatitis A virus:  
RT comparison with wild-type virus."  
RT Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).  
RL [3]  
RN [3]  
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.  
RX MEDLINE=85166289; PubMed=2984684;  
RA Baroudy B.M., Ticehurst J.R., Mele T.A., Matzel J.V. Jr.,  
RA Purcell R.H., Felstone S.M.;  
RT "Sequence analysis of hepatitis A virus cDNA coding for capsid  
RT proteins and RNA polymerase."  
RT Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).  
RL [4]  
RN [4]  
RP CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +  
CC (RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED  
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -1- CAUTION: REF. 3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT  
CC SHOWN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch).

CC -----  
 DR EMBL: M14114; AAA45475.1; -  
 DR EMBL: M14707; AAA45465.1; -  
 DR EMBL: M14707; AAA45466.1; ALT\_INIT.  
 DR EMBL: M16632; AAA45471.1; -  
 DR PIR: A23981; GNNYHM.  
 DR PIR: A24149; GNNYHM.  
 DR PDB: 1HAV; 23-DEC-96.  
 DR MEROPS: C03.005; -  
 DR InterPro: IPR004004; Calicl\_pol\_hel.  
 DR InterPro: IPR000605; RNA\_helicase.  
 DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro: IPR001205; RNA\_pol\_P3D.  
 DR InterPro: IPR007094; RNA\_pol\_PSVlr.  
 DR Pfam: PF00680; RNA\_dep\_RNA\_pol. 1.  
 DR Pfam: PF00910; RNA\_helicase. 1.  
 DR PRINTS: PR00918; CALICVIRUSNS.  
 KM Polypeptide; Coat protein; Core protein; Transferrase; RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.  
 FT CHAIN 1 23  
 FT CHAIN 24 245  
 FT CHAIN 246 491  
 FT CHAIN 492 836  
 FT CHAIN 837 980  
 FT CHAIN 981 1087  
 FT CHAIN 1088 1422  
 FT CHAIN 1423 1496  
 FT CHAIN 1497 1519  
 FT CHAIN 1520 1738  
 FT CHAIN 1739 2227  
 FT VARIANT 77 77  
 FT VARIANT 764 764  
 FT VARIANT 821 821  
 FT VARIANT 1052 1052  
 FT VARIANT 1062 1062  
 FT VARIANT 1118 1118  
 FT VARIANT 1151 1151  
 FT VARIANT 1163 1163  
 FT VARIANT 1277 1277  
 FT VARIANT 1500 1500  
 FT VARIANT 1805 1805  
 FT VARIANT 1930 1930  
 FT SEQUENCE 2227 AA; 251506 MW; 01E225E7AE8740A6 CRC64;  
 SQ

Query Match 100.0%; Score 104; DB 1; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNFPFGMDLLEIANSKD 20  
 DB 922 KVNFPFGMDLLEIANSKD 941

RESULT 4  
 POLG\_HPAVL STANDARD; PRT: 2227 AA.  
 AC POLG441;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
 DE Hepatitis A virus (strain LA).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Hepatovirus.  
 OX NCBI\_TaxID=12099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85190549; PubMed=2986127;  
 RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A., Merryweather J., van Nest G., Dina D.;

RT "Primary structure and gene organization of human hepatitis A virus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
 CC [RNA](N).  
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
 CC VP3, AND VP4.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sb-sib.ch](mailto:license@sb-sib.ch)).

CC -----  
 DR EMBL: K02990; AAA45472.1; -  
 DR PIR: A03903; GNNYHM.  
 DR MEROPS: C03.005; -  
 DR InterPro: IPR004004; Calicl\_pol\_hel.  
 DR InterPro: IPR000605; RNA\_helicase.  
 DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro: IPR001205; RNA\_pol\_P3D.  
 DR InterPro: IPR007094; RNA\_pol\_PSVlr.  
 DR Pfam: PF00680; RNA\_dep\_RNA\_pol. 1.  
 DR Pfam: PF00910; RNA\_helicase. 1.  
 DR PRINTS: PR00918; CALICVIRUSNS.  
 KM Polypeptide; Coat protein; Core protein; Transferrase; RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
 FT CHAIN 1 23  
 FT CHAIN 24 245  
 FT CHAIN 246 491  
 FT CHAIN 492 836  
 FT CHAIN 837 980  
 FT CHAIN 981 1076  
 FT CHAIN 1077 1422  
 FT CHAIN 1423 1484  
 FT CHAIN 1485 1507  
 FT CHAIN 1508 1678  
 FT CHAIN 1679 2227  
 FT SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;  
 SQ

Query Match 100.0%; Score 104; DB 1; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNFPFGMDLLEIANSKD 20  
 DB 922 KVNFPFGMDLLEIANSKD 941

RESULT 5  
 POLG\_HPAVM STANDARD; PRT: 2227 AA.  
 AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;  
 AC Q81090; Q81091; Q81092; Q81093;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
 DE Hepatitis A virus (strain MEB).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Hepatovirus.  
 OX NCBI\_TaxID=12100;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88045071; PubMed=2823500;  
 RA Paul A.V., Tada H., der Helm K., Wissel T., Kiehn R., Wimmer E., Delnhardt F.;

```
RT "The entire nucleotide sequence of the genome of human hepatitis A
RL virus (isolate MB).";
CC Virus Res. 8:153-171(1987).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
CC WHO HAD BEEN INJECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: M20273; AAA45474.1; -.
DR MEROPS: C03.005; -.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVlr.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
KM Polyprotein: Coat protein; Core protein; Transferase;
KM RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 836
FT CHAIN 837 980
FT CHAIN 981 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1496
FT CHAIN 1497 1519
FT CHAIN 1520 1738
FT CHAIN 1739 2227
SQ SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match 100.0%; Score 104; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2;je-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMLEDEETANSKD 20
922 KVNPHGMLEDEETANSKD 941
```

```
RT structure and growth in cell culture with other HAV strains.";
RL J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE-89232168; PubMed-2541023;
RA Balayan M.S., Kusov Y.Y., Andjapartidze A.G., Tsarev S.A.,
RA Sverdiolov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RA "Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses.";
RL FEBS Lett. 247:425-428(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC (RNA)(N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: D00924; BAA00766.1; -.
DR EMBL: X15461; CA33490.1; -.
DR PIR: A30470; GNNVSA.
DR MEROPS: C03.005; -.
DR InterPro: IPR004004; Calicet_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVlr.
DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PRINTS: PR00918; CALICIVIRUSNS.
KM Polyprotein: Coat protein; Core protein; Transferase;
KM RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
FT CHAIN 1 27
FT CHAIN 28 249
FT CHAIN 250 495
FT CHAIN 496 795
FT CHAIN 796 984
FT CHAIN 985 1091
FT CHAIN 1092 1426
FT CHAIN 1427 1498
FT CHAIN 1499 1521
FT CHAIN 1522 1741
FT CHAIN 1742 2230
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 97.1%; Score 101; DB 1; Length 2230;
Best Local Similarity 95.0%; Pred. No. 7;7e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMLEDEETANSKD 20
926 KVNPHGMLEDEETANSKD 945
```



OC Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatitis virus.  
OX NCBI\_TaxID=12094;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE-91162758; PubMed-1705995;  
RA Lemon S.M., Murphy P.C., Shelds P.A., Ping L.H., Feinstone S.M.,  
RT "Antigenic and genetic variation in cytopathic hepatitis A virus  
RT variants arising during persistent infection: evidence for genetic  
RT recombination."  
RL J. Virol. 65:2056-2065(1991).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC (RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VITRO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: M59810; AAA45468.1; -  
DR MEROPS: C03.005; -  
DR InterPro: IPR004004; Calic1\_pol\_hel.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_PS.  
DR InterPro: IPR001205; RNA\_pol\_PS.  
DR InterPro: IPR007094; RNA\_pol\_PSVir.  
DR Pfam: PF00680; RNA\_dep\_RNA\_pol.1.  
DR Pfam: PF00910; RNA\_helicase.1.  
DR PRINTS: PR00918; CALICIVIRUS.  
DR PolyProtein: Coat protein; Transferase;  
DR RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
KW CHAIN 1 23  
FT CHAIN 24 245 COAT PROTEIN VP4 (PIA).  
FT CHAIN 246 491 COAT PROTEIN VP2 (PIB).  
FT CHAIN 492 794 COAT PROTEIN VP3 (PIC).  
FT CHAIN 795 900 COAT PROTEIN VP1 (PID).  
FT CHAIN 901 1087 CORE PROTEIN P2A.  
FT CHAIN 1088 1422 CORE PROTEIN P2B.  
FT CHAIN 1423 1495 CORE PROTEIN P2C.  
FT CHAIN 1496 1518 PROBABLE PROTEIN P3A.  
FT CHAIN 1519 1737 PROBABLE PROTEIN P3B.  
FT CHAIN 1738 2226 PROBABLE PROTEIN P3C.  
FT CHAIN 2226 251152 MW; 6CD85A91D6B4E2BF CRC64;  
SO SEQUENCE  
Query Match 96.2%; Score 100; DB 1; Length 2226;  
Best Local Similarity 95.0%; Pred. No. 1,1e-08;  
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KVNPFHGMLELEIANSKD 20  
DB 922 KVNPFHGMLELEIANSKRE 941  
|||||  
RESULT 8  
RECD\_LACLA STANDARD: PRT: 387 AA.  
AC 001840; 09CIR3;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE RecA protein, chromosomal (Recombinase A).  
GN RECA OR L10354.  
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

OX NCBI\_TaxID=1360;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=ML3;  
RX MEDLINE-92384590; PubMed-1514816;  
RA Duval P., Ehrlich S.D., Gruss A.;  
RT "Use of degenerate primers for polymerase chain reaction cloning and  
RT sequencing of the Lactococcus lactis subsp. lactis recA gene."  
RL Appl. Environ. Microbiol. 58:2674-2678(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LI1403;  
RX MEDLINE-2135186; PubMed-11337471;  
RA Bolotin A., Wincker P., Manger S., Jallion O., Malarre K.,  
RA Weisenbach J., Ehrlich S.D., Sorokin A.;  
RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
RT lactis ssp. lactis 111403."  
RL Genome Res. 11:731-753(2001).  
CC -1- FUNCTION: CAN CATALYZE THE HYDROLYSIS OF ATP IN THE PRESENCE OF  
CC SINGLE-STRANDED DNA, THE ATP-DEPENDENT UPTAKE OF SINGLE-STRANDED  
CC DNA BY DUPLEX DNA, AND THE ATP-DEPENDENT HYBRIDIZATION OF  
CC HOMOLOGOUS SINGLE-STRANDED DNAs. IT INTERACTS WITH LEXA CAUSING  
CC ITS ACTIVATION AND LEADING TO ITS AUTOCATALYTIC CLEAVAGE.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the recA family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: M88106; AA25216.1; -  
DR EMBL: AE006272; AA04452.1; -  
DR PIR: B86669; B86669.  
DR HAMAP: MF\_00268; 1.  
DR InterPro: IPR003593; AAA\_Arpase.  
DR InterPro: IPR001553; RecA.  
DR Pfam: PF00154; recA.1.  
DR PRINTS: PR00142; RECA.  
DR ProDom: PD000229; RECA.1.  
DR SMART: SM00382; AAA.1.  
DR PROSITE: PS00321; RECA.1.1.  
DR PROSITE: PS50162; RECA.2.1.  
DR PROSITE: PS50163; RECA.3.1.  
KW DNA damage; DNA recombination; SOS response; ATP-binding; DNA-binding;  
KW Complete proteome.  
FT NP\_BIND 80 87 ATP (BY SIMILARITY).  
FT CONFLICT 288 288 R -> S (IN REF. 1).  
FT CONFLICT 298 298 D -> E (IN REF. 1).  
FT CONFLICT 332 333 E -> DV (IN REF. 1).  
FT CONFLICT 343 343 T -> A (IN REF. 1).  
FT CONFLICT 350 351 EA -> DS (IN REF. 1).  
FT CONFLICT 357 387 TSDSTKATATATKEEKKVETFEIELELED -> EEEETA  
FT FN (IN REF. 1).  
SO SEQUENCE 387 AA; 41477 MW; A8CAFC0B827BF14F CRC64;  
Query Match 45.2%; Score 47; DB 1; Length 387;  
Best Local Similarity 45.0%; Pred. No. 2.3;  
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;  
QY 1 KVNPFHGMLELEIANSKD 20  
DB 340 KVRTAHGLDEVAETTED 359  
|||||  
RESULT 9  
SCB2\_MOUSE  
ID SCB2\_MOUSE  
AC 092218; STANDARD: PRT: 404 AA.

DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Succinyl-CoA ligase [GDP-forming], beta-chain, mitochondrial precursor  
 DE (EC 6.2.1.4) (Succinyl-CoA synthetase, beta chain) (SCS-beta) (GTP-specific succinyl-CoA synthetase beta subunit) (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart.  
 RX MEDLINE=98438536; PubMed=9765291;  
 RA Johnson J.D., Mehus J.G., Tews K., Milavetz B.I., Lambeth D.O.;  
 RT "Genetic evidence for the expression of ATP- and GTP-specific  
 RT succinyl-CoA synthetases in multicellular eucaryotes.";  
 RL J. Biol. Chem. 273:27580-27586(1998).  
 CC -1- CATALYTIC ACTIVITY: GTP + succinate + CoA -> GDP + succinyl-CoA +  
 CC phosphate.  
 CC -1- PATHWAY: SUBSTRATE LEVEL PHOSPHORYLATION STEP OF THE TRICARBOXYLIC  
 CC ACID CYCLE.  
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -1- SIMILARITY: Belongs to the succinate/malate CoA ligase beta  
 CC subunit family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF058956; AAC64399.1; -  
 DR HSSP: P07460; 1SCU  
 DR MGD; MGI:1306824; Suc1g2.  
 DR InterPro: IPR003135; ATP-grasp.  
 DR InterPro: IPR005809; CoA\_lig\_beta.  
 DR InterPro: IPR005811; CoA\_ligase.  
 DR Pfam: PF02222; ATP-grasp.1.  
 DR Pfam: PF00549; ligase-CoA.1.  
 DR TIGRfams: TIGR01016; succoAbeta.1.  
 DR PROSITE; PS01217; SUCCINYL-CoA\_lig\_3; 1.  
 DR Ligase; Glycolysis; Tricarboxylic acid cycle; Mitochondrion;  
 KW TRANSIT PEPTIDE.  
 KW NON\_TER 1 1  
 FT CHAIN 10 404 SUCCINYL-CoA LIGASE [GDP-FORMING] BETA-  
 FT CHAIN  
 SQ SEQUENCE 404 AA; 43857 MW; 511757A75883E8C4 CRC64;  
 FT  
 Query Match 45.2%; Score 47; DB 1; Length 404;  
 Best Local Similarity 47.4%; Pred. No. 2.4;  
 Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
 QY 2 VNFPHGMLDELEIANSKD 20  
 Db 140 VGSFGSRSDIEVVAASPE 158  
 ID SYG\_THETH STANDARD; PRT; 505 AA.  
 AC P56206; O50551;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Glycyl-tRNA synthetase (EC 6.1.1.14) (glycine--tRNA ligase) (GLYRS).  
 GN GLYS.  
 OS Thermus thermophilus.

CC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;  
 CC Thermus.  
 OX NCBI\_TaxID=274;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND X-RAY CRYSTALLOGRAPHY (2.75 ANGSTROMS).  
 RC STRAIN=HB8 / ATCC 27634;  
 RX MEDLINE=98149692; PubMed=9490048;  
 RA Mazauric M.-H., Keth G., Logan D., Kreutzer R., Giege R., Kern D.;  
 RT "Glycyl-tRNA synthetase from Thermus thermophilus - wide structural  
 RT divergence with other prokaryotic glycyl-tRNA synthetases and  
 RT functional inter-relation with prokaryotic and eukaryotic glycylation  
 RT systems.";  
 RL Eur. J. Biochem. 251:744-757(1998).  
 CC -1- CATALYTIC ACTIVITY: ATP + glycine + tRNA(Gly) -> AMP + diphosphate  
 CC + glycyl-tRNA(Gly).  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AJ222643; CAA10903.1; -  
 DR PDB: 1AT1; 07-JUL-97.  
 DR PDB: 1B76; 28-JAN-99.  
 DR PDB: 1GGM; 28-JAN-99.  
 DR HAMAP; MF\_00253; -; 1.  
 DR InterPro: IPR004154; HGTP\_anticonodon.  
 DR InterPro: IPR002314; tRNA-synt\_2b.  
 DR InterPro: IPR002315; tRNA-synt\_gly.  
 DR InterPro: IPR006195; tRNA\_ligase\_II.  
 DR Pfam; PF03129; HGTP\_anticonodon.1.  
 DR Pfam; PF00587; tRNA-synt\_2b; 1.  
 DR PRINTS; PRO1043; TRNASYNTGLY.  
 DR TIGRfams: TIGR00389; glys\_dimeric.1.  
 DR PROSITE; PS50862; AA\_tRNA\_LIGASE\_II; 1.  
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW 3D-structure.  
 FT INIT\_MET 0 0  
 FT CONFLICT 1 1  
 FT CONFLICT 115 119 YRAME -> TEPMR (IN REF. 2).  
 FT CONFLICT 191 199 QGIVNFRN -> RASSSTST (IN REF. 2).  
 FT CONFLICT 215 219 IGKAF -> SARPS (IN REF. 2).  
 FT CONFLICT 266 266 E -> R (IN REF. 2).  
 FT CONFLICT 283 284 SS -> EL (IN REF. 2).  
 FT CONFLICT 302 303 SL -> LE (IN REF. 2).  
 FT CONFLICT 310 310 Q -> N (IN REF. 2).  
 FT HELIX 5 14  
 FT TURN 15 16  
 FT STRAND 18 20  
 FT TURN 21 22  
 FT HELIX 23 25  
 FT TURN 26 26  
 FT TURN 29 30  
 FT STRAND 32 34  
 FT TURN 36 53  
 FT HELIX 54 55  
 FT STRAND 60 64  
 FT STRAND 68 70  
 FT TURN 71 72

```

FT HELIX 73 76
FT TURN 77 77
FT HELIX 79 82
FT TURN 83 83
FT STRAND 168 168
FT STRAND 170 172
FT STRAND 179 181
FT STRAND 182 185
FT HELIX 190 195
FT TURN 196 196
FT HELIX 197 204
FT TURN 205 205
FT STRAND 210 220
FT TURN 227 227
FT HELIX 228 230
FT STRAND 234 244
FT HELIX 246 248
FT HELIX 249 266
FT TURN 267 268
FT HELIX 271 273
FT STRAND 274 278
FT TURN 281 283
FT TURN 286 287
FT TURN 290 298
FT TURN 299 300
FT STRAND 301 310
FT TURN 312 313
FT HELIX 314 319
FT TURN 323 327
FT STRAND 344 345
FT STRAND 352 353
FT STRAND 356 363
FT HELIX 364 375
FT STRAND 376 380
FT TURN 382 383
FT STRAND 386 390
FT HELIX 394 396
FT STRAND 401 405
FT HELIX 411 425
FT TURN 426 427
FT STRAND 432 434
FT HELIX 440 449
FT TURN 450 451
FT STRAND 454 458
FT HELIX 460 463
FT TURN 464 468
FT TURN 467 475
FT TURN 472 475
FT TURN 476 481
FT STRAND 482 484
FT STRAND 487 491
FT HELIX 492 503
SQ SEQUENCE 505 AA; 58081 MM; EC1C8D5388AB7570 CRC64;

Query Match 45.2%; Score 47; DB 1; Length 505;
Best Local Similarity 75.0%; Pred. No. 3;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 4 FPHGMLDEETIA 15
DB 298 FPHGSLDEGIA 309

RESULT 11
DEOC_VIRCH STANDARD: PRT: 259 AA.
AC 09KPL7: 2001 (Rel. 40, Created)
DT 16-OCt-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Deoxyribose-phosphate aldolase (EC 4.1.2.4) (Phosphodeoxyribosaldolase)
DE (Deoxyribosaldolase) (DERA).
GN DEOC OR VC2330.

```

```

OS Vibrio Cholerae.
OC Bacteria: Proteobacteria: Gammaproteobacteria: Vibrionales;
OC Vibrionaceae: Vibrio.
OX NCBI_Taxid=666;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,
RA Dodson R.J., Hart D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Yamatchevan J., Baas S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -1- CATALYTIC ACTIVITY: 2-deoxy-D-ribose 5-phosphate - D-
CC glycerinaldehyde 3-phosphate + acetaldehyde.
CC -1- PATHWAY: Nucleotide and deoxyribonucleotide catabolism.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE DEOC/FBAB FAMILY OF ALDOLASES. DEOC
CC SUBFAMILY 2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AF004305; AAF95493.1; -
DR PIR: F82087; F82087.
DR TIGR: VC2350; -
DR HAMAP: MF_00592; -; 1.
DR InterPro: IPR002915; DEOC.
DR Pfam: PF01791; DEOC; 1.
KW Lyase; Schiff base; Complete proteome.
FT BINDING 166 166 SCHIFF-BASE WITH ACETALDEHYDE (BY
FT ACT_SITE 200 200 SIMILARITY).
FT BY SIMILARITY.
SQ SEQUENCE 259 AA; 27959 MM; 2876B07AC6527C73 CRC64;

Query Match 44.2%; Score 46; DB 1; Length 259;
Best Local Similarity 52.9%; Pred. No. 2.2;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
OY 3 NFPHGMLDEETIANSK 19
DB 75 NFPHGNDDEETAVAEK 91

RESULT 12
SCB2_HUMAN STANDARD: PRT: 432 AA.
ID SCB2_HUMAN 095195; O8W0L1;
AC 095199; 095195; O8W0L1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Succinyl-CoA ligase (GDP-forming) beta-chain, mitochondrial precursor
DE (EC 6.2.1.4) (Succinyl-CoA synthetase, beta-chain) (SCS-beta) (GTP-
DE specific succinyl-CoA synthetase beta subunit) (Fragment).
GN SUCG2.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta, and Uterus;
RC MEDLINE=22388257; PubMed=12477932;

```

RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,  
 RA Krausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uscid T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.R., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smillius D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "generation and initial analysis of more than 15,000 full-length  
 human and mouse cDNA sequences.",  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [2]  
 [3]  
 CC SEQUENCE OF 29-432 FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=98438536; Pubmed=9765291;  
 RA Johnson J.D., Mehus J.G., Tewes K., Milavetz B.I., Lambeth D.O.,  
 RT "genetic evidence for the expression of ATP- and GTP-specific  
 succinyl-CoA synthetases in multicellular eucaryotes.",  
 RL J. Biol. Chem. 273:27580-27586(1998).  
 RP SEQUENCE OF 234-432 FROM N.A.  
 RA Mel G., Yu W., Gibbs R.A.,  
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: GTP + succinate + CoA = GDP + succinyl-CoA +  
 phosphate.  
 CC -1- PATHWAY: SUBSTRATE LEVEL PHOSPHORYLATION STEP OF THE TRICARBOXYLIC  
 ACID CYCLE.  
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -1- SIMILARITY: Belongs to the succinate/malate CoA ligase beta  
 subunit family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL: BC007716; AAH07716.1; -  
 DR EMBL: BC019868; AAH19868.1; -  
 DR EMBL: AF038954; AAC64397.1; -  
 DR EMBL: AF131748; AAD20032.1; -  
 DR HSSP: P07460; ISCUI.  
 DR Genew: HGNC:11450; SUCLG2.  
 DR MIM: 603922; -  
 DR GO: GO:0008325; C:succinate-CoA ligase complex (GDP-forming) . . .; NAS.  
 DR GO: GO:0004776; F:succinate-CoA ligase (GDP-forming) activity; NAS.  
 DR GO: GO:0006104; P:succinyl-CoA metabolism; NAS.  
 DR InterPro: IPR003135; ATP-grasp.  
 DR InterPro: IPR005809; CoA\_lig\_beta.  
 DR InterPro: IPR005811; CoA\_ligase.  
 DR Pfam: PF02222; ATP-grasp.1.  
 DR Pfam: PF00349; ligase-CoA.1.  
 DR TIGRfam: TIGR01016; succoabeta.1.  
 DR PROSITE: PS01217; SUCCINYL\_COA\_LIG\_3; 1.  
 DR Ligase: glycolysis; Tricarboxylic acid cycle; Mitochondrion;  
 KM Transist peptide.  
 FT NON\_TER 1 1  
 FT TRANSIT <1 37 MITOCHONDRION (BY SIMILARITY).  
 FT CHAIN 38 432 SUCCINYL-CoA LIGASE [GDP-FORMING] BETA-  
 FT CHAIN  
 FT CONFLICT 174 175 GV -> RS (IN REF. 2).

SQ SEQUENCE 432 AA; 46535 MW; EB4D025B42EA7BE0 CRC64;  
 Query Match 43.3%; Score 45; DB 1; Length 432;  
 Best Local Similarity 42.1%; Pred. No. 5.6;  
 Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 Oy 2 VNFPHGLDIEETANSKD 20  
 Db 168 VGSFGGVDFIEVAASNPE 186  
 RESULT 13  
 ID SCB2\_PIG STANDARD; PRT; 433 AA.  
 AC P53590; Q95279;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Succinyl-CoA ligase [GDP-forming] beta-chain, mitochondrial precursor  
 DE (RC 6.2.1.4) (Succinyl-CoA synthetase, beta chain) (SCS-beta) (GTP-  
 DE specific succinyl-CoA synthetase beta subunit) (Fragment).  
 GN SUCLG2.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 ON NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-41.  
 RC TISSUE=Heart;  
 RX MEDLINE=94004462; Pubmed=8401211;  
 RA Bailey D.L., Wolodko W.T., Bridger W.A.,  
 RT "Cloning, characterization, and expression of the beta subunit of pig  
 heart succinyl-CoA synthetase.",  
 RL Protein Sci. 2:1255-1262(1993).  
 RN [2]  
 RP SEQUENCE OF 1-82 FROM N.A.  
 RC TISSUE=Small Intestine;  
 RX MEDLINE=96327607; Pubmed=8672129;  
 RA Winteroe A.K., Fredholm M., Davies W.,  
 RT "Evaluation and characterization of a porcine small intestine cDNA  
 library: analysis of 839 clones.",  
 RL Mamm. Genome 7:509-517(1996).  
 CC -1- CATALYTIC ACTIVITY: GTP + succinate + CoA = GDP + succinyl-CoA +  
 phosphate.  
 CC -1- PATHWAY: SUBSTRATE LEVEL PHOSPHORYLATION STEP OF THE TRICARBOXYLIC  
 ACID CYCLE.  
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta chain.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial.  
 CC -1- SIMILARITY: Belongs to the succinate/malate CoA ligase beta  
 subunit family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL: L06944; AAA31120.1; ALN\_INIT.  
 DR EMBL: Z81187; CAB03559.1; -  
 DR PDB: 1EUC; 27-JUL-00.  
 DR PDB: 1EUD; 27-JUL-00.  
 DR InterPro: IPR003135; ATP-grasp.  
 DR InterPro: IPR005809; CoA\_lig\_beta.  
 DR InterPro: IPR005811; CoA\_ligase.  
 DR Pfam: PF02222; ATP-grasp.1.  
 DR Pfam: PF00349; ligase-CoA.1.  
 DR TIGRfam: TIGR01016; succoabeta.1.  
 DR PROSITE: PS01217; SUCCINYL\_COA\_LIG\_3; 1.  
 DR Ligase: glycolysis; Tricarboxylic acid cycle; Mitochondrion;  
 KM Transist peptide; 3D-structure.  
 FT NON\_TER 1 1

FT TRANSIT <1 38 MITOCHONDRION.  
 FT CHAIN 39 433 SUCCINYL-COA LIGASE [GDP-FORMING] BETA-  
 FT SEQUENCE 433 AA; 46803 MW; AA04B72BC1B80E24 CRC64;  
 SO  
 Query Match 43.3%; Score 45; DB 1; Length 433;  
 Best Local Similarity 42.1%; Pred. No. 5.6;  
 Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 VNFPHGMLDLEIANSKD 20  
 DB 169 VGSFGVDIEVMAASNP 187  
 RESULT 14  
 ANP1\_YEAST STANDARD; PRT; 500 AA.  
 AC P32629;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DI 01-FEB-1995 (Rel. 31, Last sequence update)  
 15-SEP-2003 (Rel. 42, Last annotation update)  
 Mannan polymerase II complex ANP1 subunit (M-Pol II subunit ANP1)  
 DE (Amanitophenyl propanediol resistance protein).  
 GN ANP1 OR GEM3 OR YEL036C OR SYCP-ORF28.  
 OS Saccharomyces cerevisiae (baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-B-6441;  
 RX MEDLINE=94016558; PubMed=8411151;  
 RA Mednick L., Sherman F.;  
 RT "The gene clusters ARC and COR on chromosomes 5 and 10, respectively,  
 of Saccharomyces cerevisiae share a common ancestry.";  
 RL J. Mol. Biol. 233:372-388(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / AB972;  
 RX Dierich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,  
 RA Arujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,  
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,  
 RA Huntcke-Smith S., Hyman R.W., Kayser A., Komp C., Laskari D., Lew H.,  
 RA Lin D., Moseley D., Nakahara K., Namath A., Norgren R., Oefner P.,  
 RA Oh C., Patel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,  
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";  
 FT Nature 387:78-81(1997).  
 RN [3]  
 RP SEQUENCE OF 1-10, SUBCELLULAR LOCATION, AND SUBUNITS.  
 RX MEDLINE=98096381; PubMed=9434768;  
 RA Hashimoto H., Yoda K.;  
 RT "Novel membrane protein complexes for protein glycosylation in the  
 yeast Golgi apparatus.";  
 RL Biochem. Biophys. Res. Commun. 241:682-686(1997).  
 RN [4]  
 RP CHARACTERIZATION.  
 RX MEDLINE=95045382; PubMed=7957057;  
 RA Chapman R.E., Munro S.;  
 RT "The functioning of the yeast Golgi apparatus requires an ER protein  
 encoded by ANP1, a member of a new family of genes affecting the  
 secretory pathway.";  
 RL EMBO J. 13:4896-4907(1994).  
 RN [5]  
 RP ACTIVITY OF M-POL II COMPLEX, SUBUNITS, AND SUBCELLULAR LOCATION.  
 RX MEDLINE=98094364; PubMed=9430634;  
 RA Jungmann J., Munro S.;  
 RT "Multi-protein complexes in the cis Golgi of Saccharomyces cerevisiae  
 with alpha-1,6-mannosyltransferase activity.";  
 RL EMBO J. 17:423-434(1998).  
 RN [6]  
 RP SUBCELLULAR LOCATION.

RX MEDLINE=20558560; PubMed=11095735;  
 RA Todorow Z., Spang A., Carmack E., Yates J., Schekman R.;  
 RT "Active recycling of yeast Golgi mannosyltransferase complexes through  
 the endoplasmic reticulum.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:13643-13648(2000)  
 CC -1- FUNCTION: INVOLVED IN THE ORGANIZATION OF THE SECRETORY PATHWAY.  
 CC REQUIRED TO MAINTAIN A FUNCTIONAL GOLGI APPARATUS.  
 CC -1- FUNCTION: The M-Pol II complex possesses alpha-1,6-  
 mannosyltransferase activity and is probably involved in the  
 elongation of the mannan backbone of N-linked glycans on cell wall  
 and periplasmic proteins.  
 CC -1- SUBUNIT: Component of the M-Pol II complex composed of ANP1, MNN9,  
 MNN10, MNN11 and HOC1.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Cis-Golgi.  
 CC Recycles between endoplasmic reticulum and Golgi.  
 CC -1- SIMILARITY: BELONGS TO THE ANP1 / MNN9 / VAN1 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; S65964; AAD13971.1; -;  
 DR EMBL; I22171; AAA34426.1; -;  
 DR EMBL; S66114; AAB28440.1; -;  
 DR EMBL; I22173; AAA34937.1; -;  
 DR EMBL; U18779; AAB65006.1; -;  
 DR PIR; S50508; S50508.  
 DR PIR; S5000762; ANP1.  
 DR GO; GO:0000136; C:mannosyltransferase complex; TAS.  
 DR GO; GO:0000009; F:alpha-1,6-mannosyltransferase activity; IDA.  
 DR GO; GO:0000032; P:cell wall mannoprotein biosynthesis; TAS.  
 DR GO; GO:0006487; P:N-linked glycosylation; IDA.  
 DR InterPro; IPR005109; ANP1.  
 DR Pfam; PF03452; ANP1; 1.  
 KW Glycoprotein; Transmembrane; Signal-anchor; Golgi stack;  
 KW Endoplasmic reticulum.  
 FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 16 27 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).  
 FT DOMAIN 28 500 LUMENAL (POTENTIAL).  
 FT DOMAIN 446 473 GLN-RICH.  
 FT CONFLICT 220 224 HHDKD -> QSGOGN (IN REF. 1).  
 FT CONFLICT 313 313 F -> L (IN REF. 1).  
 FT CONFLICT 472 500 PGCKPLDNDKKKKHKKPKVEPLDPPDRN -> RGNGLMT  
 FT TTRTRNLTAKKH (IN REF. 1).  
 SQ SEQUENCE 500 AA; 58182 MW; 845B395CE54BCD14 CRC64;  
 QY 2 VNFPHGMLDLEIANSKD 20  
 DB 106 MTPPHNLIDSLFVSDSD 124  
 Query Match 42.3%; Score 44; DB 1; Length 500;  
 Best Local Similarity 31.6%; Pred. No. 9.7;  
 Matches 6; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
 RESULT 15  
 SRG1\_HUMAN STANDARD; PRT; 98 AA.  
 AC 075711;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Scrape-responsive protein 1 precursor (SCRG-1).  
 GN SCRG1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

```

RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE-98324999; PubMed-9660755;
RA Dron M., Dandoy-Dron F., Guillo F., Bendoudjema L., Hauw J.J.,
RA Lebon P., Dormont D., Tovey M.G.;
RT "Characterization of the human analogue of a Scrapie-responsive
   gene.";
RL J. Biol. Chem. 273:18015-18018(1998).
CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
CC -1- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY IN THE CENTRAL NERVOUS
CC SYSTEM OF ADULT, BUT NOT AT ALL IN FETAL BRAIN. HIGH LEVELS OF
CC SCRG1 TRANSCRIPTS ARE ALSO OBSERVED IN TESTIS AND AORTA.
CC -1- SIMILARITY: BELONGS TO THE SCRG1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation.
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ224677; CAA12059.1; -.
DR MIM; 603163; -.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0007399; P:neurogenesis; TAS.
KW Signal.
FT SIGNAL 1 20
FT CHAIN 21 98 POTENTIAL.
FT CARBOHYD 72 72 SCRAPIE-RESPONSIVE PROTEIN 1.
SQ SEQUENCE 98 AA: 11081 MW: 11081 MW: ASFIPD40BFS401C0 CRC64;
   N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 41.38; Score 43; DB 1; Length 98;
Best Local Similarity 44.48; Pred. NO. 2.5;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;
OY 3 NFPHGMDEETIANSKD 20
DB 40 NPEGVADLTQIDYVVD 57

```

Search completed: October 1, 2003, 09:57:50  
 Job time : 8.58824 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 33.1765 Seconds  
(without alignments)  
155.564 Million cell updates/sec

Title: US-09-171-432A-44

Perfect score: 104

Sequence: 1 KVNPHGMLEIRANSKD 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_23:\*\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaea:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	251	12 Q9ENQ4	Q9enq4 hepatitis a
2	104	100.0	251	12 Q9ENQ7	Q9enq7 hepatitis a
3	104	100.0	251	12 Q9ENQ2	Q9enq2 hepatitis a
4	104	100.0	251	12 Q9ENQ2	Q9enq2 hepatitis a
5	104	100.0	251	12 Q9ENQ6	Q9enq6 hepatitis a
6	104	100.0	251	12 Q9ENQ1	Q9enq1 hepatitis a
7	104	100.0	251	12 Q9ENQ1	Q9enq1 hepatitis a
8	104	100.0	251	12 Q9ENQ1	Q9enq1 hepatitis a
9	104	100.0	251	12 Q9ENQ9	Q9enq9 hepatitis a
10	104	100.0	251	12 Q9ENQ4	Q9enq4 hepatitis a
11	104	100.0	251	12 Q9ENQ5	Q9enq5 hepatitis a
12	104	100.0	251	12 Q9ENQ7	Q9enq7 hepatitis a
13	104	100.0	251	12 Q9ENQ5	Q9enq5 hepatitis a
14	104	100.0	251	12 Q9ENQ9	Q9enq9 hepatitis a
15	104	100.0	251	12 Q9ENQ6	Q9enq6 hepatitis a
16	104	100.0	251	12 Q9ENQ5	Q9enq5 hepatitis a

17	104	100.0	251	12 Q9ENQ7	Q9enq7 hepatitis a
18	104	100.0	1124	12 Q84780	Q84780 hepatitis a
19	104	100.0	1161	12 Q05794	Q05794 hepatitis a
20	104	100.0	2216	12 Q9WMA2	Q9wma2 hepatitis a
21	104	100.0	2218	12 Q67824	Q67824 hepatitis a
22	104	100.0	2225	12 Q9DLJ2	Q9dlj2 hepatitis a
23	104	100.0	2227	12 Q9WMA3	Q9wma3 hepatitis a
24	104	100.0	2227	12 Q9WMA3	Q9wma3 hepatitis a
25	104	100.0	2227	12 Q67825	Q67825 hepatitis a
26	104	100.0	2227	12 Q9WMA1	Q9wma1 hepatitis a
27	104	100.0	2227	12 Q67826	Q67826 hepatitis a
28	104	100.0	2227	12 Q8VON6	Q8von6 hepatitis a
29	104	100.0	2227	12 Q9IFH5	Q9ifh5 hepatitis a
30	104	100.0	2227	12 Q9WMA4	Q9wma4 hepatitis a
31	98	94.2	251	12 Q9ENQ8	Q9enq8 hepatitis a
32	98	94.2	251	12 Q9ENQ3	Q9enq3 hepatitis a
33	98	94.2	251	12 Q9ENQ0	Q9enq0 hepatitis a
34	98	94.2	251	12 Q9ENQ0	Q9enq0 hepatitis a
35	98	94.2	251	12 Q9ENQ3	Q9enq3 hepatitis a
36	98	94.2	251	12 Q9ENQ6	Q9enq6 hepatitis a
37	98	94.2	251	12 Q9ENQ8	Q9enq8 hepatitis a
38	98	94.2	251	12 Q9ENQ3	Q9enq3 hepatitis a
39	98	94.2	251	12 Q9ENQ8	Q9enq8 hepatitis a
40	98	94.2	251	12 Q9ENQ0	Q9enq0 hepatitis a
41	98	94.2	2227	12 Q8QV03	Q8qv03 hepatitis a
42	97	93.3	184	12 Q87092	Q87092 simian hepa
43	97	93.3	2225	12 Q9DMR1	Q9dmr1 hepatitis a
44	95	91.3	2218	12 Q67817	Q67817 hepatitis a
45	94	90.4	251	12 Q9ENQ2	Q9enq2 hepatitis a

## ALIGNMENTS

RESULT 1  
Q9ENQ4 PRELIMINARY; PRT; 251 AA.  
AC Q9ENQ4: 01-MAR-2001 (TREMblrel. 16, Created)  
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
DE Polyprotein (fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A201;  
RA Fujiwara K.;  
RT "Hepatitis A virus."  
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AB047659; BAB12167.1; .  
FT NON-TER  
FT NON-TER  
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. NO. 2.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNPHGMLEIRANSKD 20  
DB 86 KVNPHGMLEIRANSKD 105

## RESULT 2

Q9ENQ7 PRELIMINARY; PRT; 251 AA..  
AC Q9ENQ7: 01-MAR-2001 (TREMblrel. 16, Created)  
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)

DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A161;  
RA Fujiwara K.;  
RT "Hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047656; BABI2164.1; -.  
FT NON\_TER 1  
FT SEQUENCE 251 AA; 28766 MW; A0EC02E3609C4D5B CRC64;  
SQ

Query Match 100.0%; Score 104; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMDLEETANSKD 20  
|||||  
86 KVNPHGMDLEETANSKD 105

RESULT 3  
Q9ENN2 PRELIMINARY; PRT; 251 AA.  
AC Q9ENN2;  
DT 01-MAR-2001 (TRENBLREL. 16, Created)  
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)  
DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A9;  
RA Fujiwara K.;  
RT "Hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047681; BABI2189.1; -.  
FT NON\_TER 1  
FT SEQUENCE 251 AA; 28720 MW; C334248282F19CA CRC64;  
SQ

Query Match 100.0%; Score 104; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMDLEETANSKD 20  
|||||  
86 KVNPHGMDLEETANSKD 105

RESULT 4  
Q9ENP2 PRELIMINARY; PRT; 251 AA.  
AC Q9ENP2;  
DT 01-MAR-2001 (TRENBLREL. 16, Created)  
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)  
DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A5;  
RA Fujiwara K.;

RT "Hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047671; BABI2179.1; -.  
FT NON\_TER 1  
FT SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;  
SQ

Query Match 100.0%; Score 104; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMDLEETANSKD 20  
|||||  
86 KVNPHGMDLEETANSKD 105

RESULT 5  
Q9EN06 PRELIMINARY; PRT; 251 AA.  
AC Q9EN06;  
DT 01-MAR-2001 (TRENBLREL. 16, Created)  
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)  
DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A162;  
RA Fujiwara K.;  
RT "Hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047657; BABI2165.1; -.  
FT NON\_TER 1  
FT SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6D3B CRC64;  
SQ

Query Match 100.0%; Score 104; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPHGMDLEETANSKD 20  
|||||  
86 KVNPHGMDLEETANSKD 105

RESULT 6  
Q9ENR1 PRELIMINARY; PRT; 251 AA.  
AC Q9ENR1;  
DT 01-MAR-2001 (TRENBLREL. 16, Created)  
DT 01-MAR-2001 (TRENBLREL. 16, Last sequence update)  
DT 01-MAR-2001 (TRENBLREL. 16, Last annotation update)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A1;  
RA Fujiwara K.;  
RT "Hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047652; BABI2160.1; -.  
FT NON\_TER 1  
FT SEQUENCE 251 AA; 28749 MW; 58A520D873893445 CRC64;  
SQ

Query Match 100.0%; Score 104; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 2.4e-09;



Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KYNPFGMDLEIANSKD 20  
|||||  
Db 86 KYNPFGMDLEIANSKD 105

## RESULT 7

ID O9ENP1 PRELIMINARY; PRT; 251 AA.  
AC O9ENP1:  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A503;  
RA Fujiwara K.;  
RT "hepatitis A virus."  
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047672; BAB12180.1; -.  
FT NON\_TER 1 251  
SQ SEQUENCE 251 AA; 28663 MW; C7EA66BDD19A1619 CRC64;

## Query Match

Best Local Similarity 100.0%; Score 104; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KYNPFGMDLEIANSKD 20  
|||||  
Db 86 KYNPFGMDLEIANSKD 105

## RESULT 8

ID O9EN01 PRELIMINARY; PRT; 251 AA.  
AC O9EN01:  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A206;  
RA Fujiwara K.;  
RT "hepatitis A virus."  
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047662; BAB12170.1; -.  
FT NON\_TER 1 251  
SQ SEQUENCE 251 AA; 28699 MW; 8EADAE7E2754C37 CRC64;

## Query Match

Best Local Similarity 100.0%; Score 104; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KYNPFGMDLEIANSKD 20  
|||||  
Db 86 KYNPFGMDLEIANSKD 105

## RESULT 9

OY 1 KYNPFGMDLEIANSKD 20  
|||||  
Db 86 KYNPFGMDLEIANSKD 105

ID O9EN09 PRELIMINARY; PRT; 251 AA.  
AC O9EN09:  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A159;  
RA Fujiwara K.;  
RT "hepatitis A virus."  
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047654; BAB12162.1; -.  
FT NON\_TER 1 251  
SQ SEQUENCE 251 AA; 28720 MW; C342482882F19CA CRC64;

## Query Match

Best Local Similarity 100.0%; Score 104; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KYNPFGMDLEIANSKD 20  
|||||  
Db 86 KYNPFGMDLEIANSKD 105

## RESULT 10

ID O9EN04 PRELIMINARY; PRT; 251 AA.  
AC O9EN04:  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A77;  
RA Fujiwara K.;  
RT "hepatitis A virus."  
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047679; BAB12187.1; -.  
FT NON\_TER 1 251  
SQ SEQUENCE 251 AA; 28734 MW; 3B96B782882F19D9 CRC64;

## Query Match

Best Local Similarity 100.0%; Score 104; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 2.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KYNPFGMDLEIANSKD 20  
|||||  
Db 86 KYNPFGMDLEIANSKD 105

## RESULT 11

ID O9ENP5 PRELIMINARY; PRT; 251 AA.  
AC O9ENP5:  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.

```

OX NCB1_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A407;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047668; BAB12176.1; -.
FT NON_TER
FT NON_TER
SO SEQUENCE 251 AA; 28614 MW; 8334EF19C757A6D CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. NO. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPPHGMLDLEETANSKD 20
DB 86 KVNPPHGMLDLEETANSKD 105
OY [1]
DB [1]
OYENP7 PRELIMINARY; PRT; 251 AA.
AC O9ENP7;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A306;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047666; BAB12174.1; -.
FT NON_TER
FT NON_TER
SO SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 104; DB 12; Length 251;
Best Local Similarity 100.0%; Pred. NO. 2.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KVNPPHGMLDLEETANSKD 20
86 KVNPPHGMLDLEETANSKD 105

RESULT 13
O9ENOS PRELIMINARY; PRT; 251 AA.
AC O9ENOS;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A20;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047658; BAB12166.1; -.
FT NON_TER
FT NON_TER

```

FT	NON_TER	251	251	MM	C3342482882F19CA	CRC64
SO	SEQUENCE	251 AA	28720 MW			
Query Match						
Best Local Similarity		100.0%	Score 104	DB 12	Length 251	
Matches 20		Conservative 0	Mismatches 0	Indels 0	Gaps 0	
OY	1	KVNFPHGMLDEETAA	SKD 20			
Db 86 KVNFPHGMLDEETAA						
SKD 105						
RESULT 14						
OY	Q9ENP9	PRELIMINARY	PRT	251 AA		
AC	Q9ENP9					
DT	01-MAR-2001	(TREMBLrel. 16, Created)				
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)				
DT	01-MAR-2001	(TREMBLrel. 16, Last annotation update)				
DE	Polyprotein (Fragment)					
OS	Hepatitis A virus					
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus					
CC	NCBI_TaxID=12092					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN-A303					
RA	Fujiwara K.;					
RT	"Hepatitis A virus";					
RL	Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.					
EMBL	AB047664; BAB12172.1; -					
FT	NON_TER	1				
FT	NON_TER	251				
SO	SEQUENCE	251 AA	28752 MW	7215A28AD2CA5C1A	CRC64	
Query Match						
Best Local Similarity		100.0%	Score 104	DB 12	Length 251	
Matches 20		Conservative 0	Mismatches 0	Indels 0	Gaps 0	
OY	1	KVNFPHGMLDEETAA	SKD 20			
Db 86 KVNFPHGMLDEETAA						
SKD 105						
RESULT 15						
OY	Q9ENN6	PRELIMINARY	PRT	251 AA		
AC	Q9ENN6					
DT	01-MAR-2001	(TREMBLrel. 16, Created)				
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)				
DT	01-MAR-2001	(TREMBLrel. 16, Last annotation update)				
DE	Polyprotein (Fragment)					
OS	Hepatitis A virus					
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus					
CC	NCBI_TaxID=12092					
OX	[1]					
RN	SEQUENCE FROM N.A.					
RC	STRAIN-A713					
RA	Fujiwara K.;					
RT	"Hepatitis A virus";					
RL	Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.					
EMBL	AB047677; BAB12185.1; -					
FT	NON_TER	1				
FT	NON_TER	251				
SO	SEQUENCE	251 AA	28720 MW	C3342482882F19CA	CRC64	
Query Match						
Best Local Similarity		100.0%	Score 104	DB 12	Length 251	
Matches 20		Conservative 0	Mismatches 0	Indels 0	Gaps 0	
OY	1	KVNFPHGMLDEETAA	SKD 20			
Db 86 KVNFPHGMLDEETAA						
SKD 105						

Thu Oct 2 09:13:58 2003

us-09-171-432a-44.rpt

Page 5

Db 86 KYNPFGMIDLEIANSKD 105

Search completed: October 1, 2003, 10:02:40  
Job time : 35.1765.secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 : Search time 12.7059 seconds  
(without alignments)  
66.600 Million cell updates/sec

Title: US-09-171-432a-44  
Perfect score: 104  
Sequence: 1 KVNFGMDLEIANSKD 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues  
Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Dackfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	104	100.0	2227	3	US-08-475-886-2
2	104	100.0	2227	3	US-08-475-886-4
3	104	100.0	2227	3	US-08-475-886-6
4	104	100.0	2227	3	US-08-397-232-2
5	104	100.0	2227	3	US-08-397-232-4
6	104	100.0	2227	3	US-09-171-387-2
7	104	100.0	2227	4	US-09-653-499-2
8	104	100.0	2227	4	US-09-653-499-4
9	104	100.0	2227	4	US-09-653-499-6
10	104	100.0	2227	4	US-09-653-499-6
11	45	43.3	417	2	US-09-099-677A-6
12	45	43.3	417	2	US-09-261-471-6
13	45	43.3	432	2	US-09-099-677A-3
14	43	41.3	98	3	US-09-261-471-3
15	43	41.3	98	3	US-09-181-487-2
16	41.5	39.9	331	4	US-09-227-357-219
17	40	38.5	589	4	US-09-634-238-247
18	40	38.5	589	4	US-08-939-309-8
19	40	38.5	589	4	US-09-849-180-8
20	39	37.5	226	4	US-09-356-643B-2
21	39	37.5	314	4	US-09-198-452A-210
22	39	37.5	579	5	US-09-252-991A-27913
23	39	37.5	579	5	US-08-126-564A-31
24	39	37.5	609	4	PCT-US94-09143-31
25	39	37.5	609	4	US-09-115-475-16
26	39	37.5	609	4	US-09-115-475-19
27	39	37.5	609	4	US-09-115-475-22

28	39	37.5	609	4	US-09-115-475-28	Sequence 28, Appl
29	39	37.5	609	4	US-09-115-475-31	Sequence 31, Appl
30	39	37.5	1066	4	US-09-004-838-24	Sequence 24, Appl
31	39	37.5	1323	4	US-09-004-838-90	Sequence 90, Appl
32	39	37.5	1604	4	US-09-004-838-95	Sequence 95, Appl
33	39	37.5	1805	4	US-09-004-838-92	Sequence 92, Appl
34	39	37.5	1817	4	US-09-004-838-125	Sequence 125, Appl
35	38	36.5	181	4	US-09-134-001C-3897	Sequence 3897, Ap
36	38	36.5	287	4	US-09-328-352-7035	Sequence 20036, A
37	38	36.5	375	4	US-09-252-991A-20086	Sequence 692, App
38	38	36.5	541	4	US-09-198-452A-692	Sequence 26363, A
39	38	36.5	1238	4	US-09-252-991A-9	Sequence 9, Appl1
40	38	36.5	1250	3	US-08-938-291A-9	Sequence 9, Appl1
41	38	36.5	1250	4	US-09-589-619-9	Sequence 2, Appl1
42	38	36.5	1367	2	US-08-249-687C-2	Sequence 2, Appl1
43	38	36.5	1367	2	US-08-625-819-2	Sequence 2, Appl1
44	38	36.5	1367	3	US-08-746-559A-2	Sequence 18, Appl
45	38	36.5	1367	4	US-08-864-641B-18	

ALIGNMENTS

RESULT 1  
US-08-475-886-2  
Sequence 2, Application US/08475886A  
Patent No. 6113912  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PORCELL, ROBERT H  
APPLICANT: D'HOUDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262US2  
CURRENT APPLICATION NUMBER: US/08/475,886A  
EARLIER FILING DATE: 1995-06-07  
EARLIER APPLICATION NUMBER: 07/947,338  
EARLIER FILING DATE: 1992-09-18  
EARLIER APPLICATION NUMBER: 08/397,232  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-08-475-886-2  
Query Match  
Best local Similarity 100.0%; Score 104; DB 3; Length 2227;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KVNFGMDLEIANSKD 20  
DB 922 KVNFGMDLEIANSKD 941  
US-08-475-886-4  
Sequence 4, Application US/08475886A  
Patent No. 6113912  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PORCELL, ROBERT H  
APPLICANT: D'HOUDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262US2  
CURRENT APPLICATION NUMBER: US/08/475,886A  
EARLIER FILING DATE: 1995-06-07  
EARLIER APPLICATION NUMBER: 07/947,338  
EARLIER FILING DATE: 1992-09-18  
EARLIER APPLICATION NUMBER: 08/397,232

```
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4
```

```
Query Match          100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  KVNPHGMLDEETIAANSKD 20
Db      922  KVNPHGMLDEETIAANSKD 941
```

```
RESULT 3
US-08-475-886-6
Sequence 6, Application US/08475886A
Patent No. 6113912
```

```
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6
```

```
Query Match          100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  KVNPHGMLDEETIAANSKD 20
Db      922  KVNPHGMLDEETIAANSKD 941
```

```
RESULT 4
US-08-397-232-2
Sequence 2, Application US/08397232A
Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
```

```
; TYPE: PRF
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2
```

```
Query Match          100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  KVNPHGMLDEETIAANSKD 20
Db      922  KVNPHGMLDEETIAANSKD 941
```

```
RESULT 5
US-08-397-232-4
Sequence 4, Application US/08397232A
Patent No. 6180110
```

```
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; CURRENT FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4
```

```
Query Match          100.0%; Score 104; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8.4e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1  KVNPHGMLDEETIAANSKD 20
Db      922  KVNPHGMLDEETIAANSKD 941
```

```
RESULT 6
US-09-171-387-2
Sequence 2, Application US/09171387
Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA.
```

```
; EMERSON, SUZANNE, U.;
; PURCELL, ROBERT, H.
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSER: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999
```

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/06506  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US60/015,642  
FILING DATE: 19-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: William S. Feller  
REGISTRATION NUMBER: 26,728  
REFERENCE/DOCKET NUMBER: 2026-4229US1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 2  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2227 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
09-171-387-2

Query Match 100.0%; Score 104; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 8.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPFHGMJLDEEIANSKD 20  
DB 922 KVNPFHGMJLDEEIANSKD 941

RESULT 7  
US-09-653-499-2  
Sequence 2, Application US/09653499  
Patent No. 6423318  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PORCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262US2  
CURRENT APPLICATION NUMBER: US/09/653,499  
CURRENT FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: 08/475,886  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-09-653-499-2

Query Match 100.0%; Score 104; DB 4; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 8.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPFHGMJLDEEIANSKD 20  
DB 922 KVNPFHGMJLDEEIANSKD 941

RESULT 8  
US-09-653-499-4  
Sequence 4, Application US/09653499  
Patent No. 6423318  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U

APPLICANT: PORCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262US2  
CURRENT APPLICATION NUMBER: US/09/653,499  
CURRENT FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: 08/475,886  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
US-09-653-499-4

Query Match 100.0%; Score 104; DB 4; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 8.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPFHGMJLDEEIANSKD 20  
DB 922 KVNPFHGMJLDEEIANSKD 941

RESULT 9  
US-09-653-499-6  
Sequence 6, Application US/09653499  
Patent No. 6423318  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PORCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262US2  
CURRENT APPLICATION NUMBER: US/09/653,499  
CURRENT FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: 08/475,886  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-09-653-499-6

Query Match 100.0%; Score 104; DB 4; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 8.4e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPFHGMJLDEEIANSKD 20  
DB 922 KVNPFHGMJLDEEIANSKD 941

RESULT 10  
US-09-099-677A-6  
Sequence 6, Application US/09099677A  
Patent No. 5965369  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Lai, Preeti  
APPLICANT: Corley, Neil C.

```

1 APPLICANT: Patterson, Chandra
2 TITLE OF INVENTION: HUMAN SUCCINYL-COENZYME A SYNTHETASE HOLOENZYME
3 NUMBER OF SEQUENCES: 6
4 CORRESPONDENCE ADDRESS:
5 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
6 STREET: 3174 PORTER DRIVE
7 CITY: PALO ALTO
8 STATE: CALIFORNIA
9 COUNTRY: USA
10 ZIP: 94304
11
12 COMPUTER READABLE FORM:
13 MEDIUM TYPE: Floppy disk
14 COMPUTER: IBM PC compatible
15 OPERATING SYSTEM: PC-DOS/MS-DOS
16 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
17 CURRENT APPLICATION DATA:
18 APPLICATION NUMBER: US/09/099,677A
19 FILING DATE: June 18, 1998
20 CLASSIFICATION: 435
21
22 ATTORNEY/AGENT INFORMATION:
23 NAME: CERRONE, MICHAEL C.
24 REGISTRATION NUMBER: 39,132
25 REFERENCE/DOCKET NUMBER: PR-0545 US
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: (650) 855-0555
28 TELEFAX: (650) 845-4166
29 INFORMATION FOR SEQ ID NO: 6:
30 SEQUENCE CHARACTERISTICS:
31 LENGTH: 417 amino acids
32 TYPE: amino acid
33 STRANDEDNESS: single
34 TOPOLOGY: linear
35 IMMEDIATE SOURCE:
36 LIBRARY: GenBank
37 CLONE: g164669
38
39 US-09-099-677A-6
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
96
```

```

1 APPLICATION NUMBER: 09/099,677
2 FILING DATE:
3 ATTORNEY/AGENT INFORMATION:
4 NAME: CERRONE, MICHAEL C.
5 REGISTRATION NUMBER: 39,132
6 REFERENCE/DOCKET NUMBER: PF-0545 US
7 TELECOMMUNICATION INFORMATION:
8 TELEPHONE: (650) 855-0555
9 TELEFAX: (650) 845-4166
10 INFORMATION FOR SEQ ID NO: 6:
11 SEQUENCE CHARACTERISTICS:
12 LENGTH: 417 amino acids
13 TYPE: amino acid
14 STRANDEDNESS: single
15 TOPOLOGY: linear
16 IMMEDIATE SOURCE:
17 LIBRARY: Genbank
18 CLONE: g164669
19 US-09-261-471-6
20
21 Query Match 43.3% Score 45; DB 3; Length 417;
22 Best Local Similarity 42.1% Pred. No. 11;
23 Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
24
25 QY 2 VNPPGMDLEIRIANSKD 20
26 | | | : | : | : | : | :
27 Db 153 VGSFGVDIEVAANSNE 171
28
29 RESULT 12
30 US-09-099-677A-3
31 Sequence 3, Application US/09099677A
32 Patent No. 5965369
33 GENERAL INFORMATION:
34 APPLICANT: Bandman, Olga
35 APPLICANT: Lal, Preethi
36 APPLICANT: Corley, Neil C.
37 APPLICANT: Patterson, Chandra
38 TITLE OF INVENTION: HUMAN SUCCINYL-COENZYME A SYNTHETASE HOLOENZYME
39 NUMBER OF SEQUENCES: 6
40 CORRESPONDENCE ADDRESS:
41 ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
42 STREET: 3174 PORTER DRIVE
43 CITY: PALO ALTO
44 STATE: CALIFORNIA
45 COUNTRY: USA
46 ZIP: 94304
47
48 COMPUTER READABLE FORM:
49 MEDIUM TYPE: Floppy disk
50 COMPUTER: IBM PC compatible
51 OPERATING SYSTEM: PC-DOS/MS-DOS
52 SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
53 CURRENT APPLICATION DATA:
54 APPLICATION NUMBER: US/09/099,677A
55 FILING DATE: June 18, 1998
56 CLASSIFICATION: 435
57
58 ATTORNEY/AGENT INFORMATION:
59 NAME: CERRONE, MICHAEL C.
60 REGISTRATION NUMBER: 39,132
61 REFERENCE/DOCKET NUMBER: PF-0545 US
62 TELECOMMUNICATION INFORMATION:
63 TELEPHONE: (650) 855-0555
64 TELEFAX: (650) 845-4166
65 INFORMATION FOR SEQ ID NO: 3:
66 SEQUENCE CHARACTERISTICS:
67 LENGTH: 432 amino acids
68 TYPE: amino acid
69 STRANDEDNESS: single
70 TOPOLOGY: linear
71 IMMEDIATE SOURCE:
72 LIBRARY: PROSBPT06
73 CLONE: 3273853
74
75 US-09-099-677A-3

```



Query Match 42.3%; Score 44; DB 2; Length 432;  
Best Local Similarity 42.1%; Pred. No. 18;  
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNFPHGMLEETIAANSKD 20  
DB 168 VGOPOGVDIEEVAASNP 186

## RESULT 13

US-09-261-471-3  
Sequence 3, Application US/09261471  
Patent No. 6025123  
GENERAL INFORMATION:  
APPLICANT: Bandman, Olga  
APPLICANT: Lal, Preeti  
APPLICANT: Corley, Neil C.  
TITLE OF INVENTION: HUMAN SUCCINYL-COA SYNTHETASE HOLOENZYME  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/261,471  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/099,677  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: CERONE, MICHAEL C.  
REGISTRATION NUMBER: 39,132  
REFERENCE/DOCKET NUMBER: PF-0545 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ. ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 432 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: PROSBPT06  
CLONE: 3273853  
US-09-261-471-3

Query Match 42.3%; Score 44; DB 3; Length 432;  
Best Local Similarity 42.1%; Pred. No. 18;  
Matches 8; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 2 VNFPHGMLEETIAANSKD 20  
DB 168 VGOPOGVDIEEVAASNP 186

RESULT 14  
US-09-181-487-2  
Sequence 2, Application US/09181487  
Patent No. 6165752  
GENERAL INFORMATION:  
APPLICANT: DUCKWORTH, DAVID MALCOLM  
TITLE OF INVENTION: NOVEL COMPOUNDS

NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ratner & Prestia  
STREET: P.O. Box 980  
CITY: Valley Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible

OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/181,487  
FILING DATE: 28-OCT-1998

CLASSIFICATION:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 9807082.4  
FILING DATE: 01-APR-1998

APPLICATION NUMBER: 9815489.1  
FILING DATE: 16-JUL-1998

ATTORNEY/AGENT INFORMATION:  
NAME: Prestia, Paul P.

REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: GP-30094

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0700

TELEX: 846169

INFORMATION FOR SEQ. ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 98 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-09-181-487-2

Query Match 41.3%; Score 43; DB 3; Length 98;  
Best Local Similarity 44.4%; Pred. No. 4.7;  
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 NFPHGMLEETIAANSKD 20  
DB 40 NLEPGVADLTQIDVNVQD 57

## RESULT 15

US-09-227-357-219  
Sequence 219, Application US/09227357  
Patent No. 6342581  
GENERAL INFORMATION:  
APPLICANT: Fischer et al.  
TITLE OF INVENTION: 123 Human Secreted Proteins  
FILE REFERENCE: P2010P1  
CURRENT APPLICATION NUMBER: US/09/227,357  
CURRENT FILING DATE: 1999-01-08  
EARLIER APPLICATION NUMBER: PCT/US98/13684  
EARLIER FILING DATE: 1998-07-07  
EARLIER APPLICATION NUMBER: 60/051,926  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,793  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,925  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,929  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,803  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/052,732  
EARLIER FILING DATE: 1997-07-08  
EARLIER APPLICATION NUMBER: 60/051,931

Job time : 13.7059 secs

```

; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 219
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (99)
; OTHER INFORMATION: Xaa equals stop translation
US-09-227-357-219

```

```

Query Match      41.3%; Score 43; DB 4; Length 99;
Best Local Similarity 44.4%; Pred. No. 4.7;
Matches      8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```

```

QY      3 NPPHGMLEIEIANSKD 20
      | | | | | | | | | |
Db      40 NLEPGVADLTQIDVNVOD 57

```

Search completed: October 1, 2003, 10:06:31

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 : Search time 24.1176 Seconds  
(without alignments)  
131.201 Million cell updates/sec

Title: US-09-171-432a-44  
Perfect score: 104  
Sequence: 1 KVNFPFGMDLEETIANSKD 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues  
Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PC7\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PC7US\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	100.0	2227	10 US-09-929-955-12	Sequence 12, Appl
2	104	100.0	2227	14 US-10-104-966-12	Sequence 12, Appl
3	104	100.0	2227	14 US-10-135-988-2	Sequence 2, Appl1
4	104	100.0	2227	14 US-10-135-988-4	Sequence 4, Appl1
5	104	100.0	2227	14 US-10-135-988-6	Sequence 6, Appl1
6	98	94.2	352	15 US-10-272-459-45	Sequence 45, Appl
7	98	94.2	980	15 US-10-272-459-41	Sequence 41, Appl
8	43	41.3	98	10 US-09-978-255A-616	Sequence 616, App
9	43	41.3	98	10 US-09-978-657-616	Sequence 616, App
10	43	41.3	98	10 US-09-978-192A-616	Sequence 616, App
11	43	41.3	98	10 US-09-999-832A-616	Sequence 616, App
12	43	41.3	98	11 US-09-978-189-616	Sequence 616, App
13	43	41.3	98	11 US-09-978-608A-616	Sequence 616, App
14	43	41.3	98	11 US-09-978-585A-616	Sequence 616, App
15	43	41.3	98	11 US-09-978-191A-616	Sequence 616, App

16	43	41.3	98	11 US-09-978-403A-616	Sequence 616, App
17	43	41.3	98	11 US-09-978-564A-616	Sequence 616, App
18	43	41.3	98	11 US-09-999-833A-616	Sequence 616, App
19	43	41.3	98	11 US-09-981-915A-616	Sequence 616, App
20	43	41.3	98	11 US-09-978-824-616	Sequence 616, App
21	43	41.3	98	11 US-09-918-585A-616	Sequence 616, App
22	43	41.3	98	11 US-09-978-423A-616	Sequence 616, App
23	43	41.3	98	11 US-09-978-193A-616	Sequence 616, App
24	43	41.3	98	11 US-09-999-830A-616	Sequence 616, App
25	43	41.3	98	11 US-09-978-757A-616	Sequence 616, App
26	43	41.3	98	11 US-09-978-187B-616	Sequence 616, App
27	43	41.3	98	11 US-09-978-643A-616	Sequence 616, App
28	43	41.3	98	12 US-09-978-375A-616	Sequence 616, App
29	43	41.3	98	12 US-09-978-188A-616	Sequence 616, App
30	43	41.3	98	12 US-09-978-189A-616	Sequence 616, App
31	43	41.3	98	12 US-10-143-031A-616	Sequence 616, App
32	43	41.3	98	12 US-10-002-967A-616	Sequence 616, App
33	43	41.3	98	12 US-10-017-083A-616	Sequence 616, App
34	43	41.3	98	12 US-10-143-030A-616	Sequence 616, App
35	43	41.3	98	12 US-10-216-163-106	Sequence 106, App
36	43	41.3	98	12 US-10-145-128A-616	Sequence 616, App
37	43	41.3	98	12 US-10-017-191A-616	Sequence 616, App
38	43	41.3	98	12 US-10-143-028A-616	Sequence 616, App
39	43	41.3	98	12 US-10-143-029A-616	Sequence 616, App
40	43	41.3	98	12 US-10-145-089A-616	Sequence 616, App
41	43	41.3	98	15 US-10-227-884-106	Sequence 106, App
42	43	41.3	98	15 US-10-230-163-106	Sequence 106, App
43	43	41.3	98	15 US-10-230-338-106	Sequence 106, App
44	43	41.3	98	15 US-10-218-631-106	Sequence 106, App
45	43	41.3	98	15 US-10-017-081A-616	Sequence 616, App

ALIGNMENTS

RESULT 1  
US-09-929-955-12  
; Sequence 12, Application US/09929955  
; Patent No. US20020136740A1  
; GENERAL INFORMATION:  
; APPLICANT: Matti Sallberg  
; APPLICANT: Catharina Hultgren  
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND  
; FILE REFERENCE: TRIPEP.23AUS2  
; CURRENT APPLICATION NUMBER: US/09/929,955  
; CURRENT FILING DATE: 2001-08-15  
; PRIOR APPLICATION NUMBER: 09/705,547  
; PRIOR FILING DATE: 2000-11-03  
; PRIOR APPLICATION NUMBER: 60/229,175  
; PRIOR FILING DATE: 2000-08-29  
; PRIOR APPLICATION NUMBER: 60/225,767  
; PRIOR FILING DATE: 2000-08-17  
; NUMBER OF SEQ ID NOS: 49  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 2227  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Hepatitis A virus sequence  
US-09-929-955-12

Query Match 100.0%: Score 104; DB 10;  
Best Local Similarity 100.0%: Pred. No. 1.7e+08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KVNFPFGMDLEETIANSKD 20  
DB 922 KVNFPFGMDLEETIANSKD 941

RESULT 2

US-10-104-966-12  
; Sequence 12, Application US/10104966  
; Publication No. US20020155124A1  
; GENERAL INFORMATION:  
; APPLICANT: Matti Salberg  
; APPLICANT: Catharina Hultgren  
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: TRIPEP.23AUSC1  
; CURRENT APPLICATION NUMBER: US/10/104,966  
; PRIOR FILING DATE: 2002-03-22  
; PRIOR APPLICATION NUMBER: 09/705,547  
; PRIOR FILING DATE: 2000-11-03  
; PRIOR APPLICATION NUMBER: 60/229,175  
; PRIOR FILING DATE: 2000-08-29  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 2227  
; TYPE: PRT  
FEATURE:  
OTHER INFORMATION: Hepatitis A virus sequence  
US-10-104-966-12

Query Match 100.0%; Score 104; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1,7e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMLDEETIANSKD 20  
Db 922 KVNPHGMLDEETIANSKD 941

RESULT 3  
; Sequence 2, Application US/10135988  
; Publication No. US20020176869A1  
; GENERAL INFORMATION:  
; APPLICANT: FUNKHOUSER, ANN W  
; APPLICANT: EMERSON, SUZANNE U  
; APPLICANT: PURCELL, ROBERT H  
; APPLICANT: D'HONDT, ERIC  
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
; FILE REFERENCE: 20264262US3  
; CURRENT APPLICATION NUMBER: US/10/135,988  
; PRIOR FILING DATE: 2002-04-29  
; PRIOR APPLICATION NUMBER: 07/947,338  
; PRIOR FILING DATE: 1992-09-18  
; PRIOR APPLICATION NUMBER: 08/397,232  
; PRIOR FILING DATE: 1995-03-10  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 2227  
; TYPE: PRT  
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-10-135-988-2

Query Match 100.0%; Score 104; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1,7e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMLDEETIANSKD 20  
Db 922 KVNPHGMLDEETIANSKD 941

RESULT 4  
US-10-135-988-4  
; Sequence 4, Application US/10135988  
; Publication No. US20020176869A1  
; GENERAL INFORMATION:

; APPLICANT: FUNKHOUSER, ANN W  
; APPLICANT: EMERSON, SUZANNE U  
; APPLICANT: PURCELL, ROBERT H  
; APPLICANT: D'HONDT, ERIC  
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
; FILE REFERENCE: 20264262US3  
; CURRENT APPLICATION NUMBER: US/10/135,988  
; PRIOR FILING DATE: 2002-04-29  
; PRIOR APPLICATION NUMBER: 07/947,338  
; PRIOR FILING DATE: 1992-09-18  
; PRIOR APPLICATION NUMBER: 08/397,232  
; PRIOR FILING DATE: 1995-03-10  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 2227  
; TYPE: PRT  
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
US-10-135-988-4

Query Match 100.0%; Score 104; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1,7e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMLDEETIANSKD 20  
Db 922 KVNPHGMLDEETIANSKD 941

RESULT 5  
US-10-135-988-6  
; Sequence 6, Application US/10135988  
; Publication No. US20020176869A1  
; GENERAL INFORMATION:  
; APPLICANT: FUNKHOUSER, ANN W  
; APPLICANT: EMERSON, SUZANNE U  
; APPLICANT: PURCELL, ROBERT H  
; APPLICANT: D'HONDT, ERIC  
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
; FILE REFERENCE: 20264262US3  
; CURRENT APPLICATION NUMBER: US/10/135,988  
; PRIOR FILING DATE: 2002-04-29  
; PRIOR APPLICATION NUMBER: 07/947,338  
; PRIOR FILING DATE: 1992-09-18  
; PRIOR APPLICATION NUMBER: 08/397,232  
; PRIOR FILING DATE: 1995-03-10  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 2227  
; TYPE: PRT  
; ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-10-135-988-6

Query Match 100.0%; Score 104; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1,7e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KVNPHGMLDEETIANSKD 20  
Db 922 KVNPHGMLDEETIANSKD 941

RESULT 6  
US-10-272-459-45  
; Sequence 45, Application US/10272459  
; Publication No. US20030124517A1  
; GENERAL INFORMATION:  
; APPLICANT: PICHUANTES, Sergio  
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT  
; FILE REFERENCE: Ppl7955.002 / 2301-17955  
; CURRENT APPLICATION NUMBER: US/10/272,459

```

CURRENT FILING DATE: 2002-10-15
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 45
LENGTH: 352
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-45

Query Match          94.2%; Score 98; DB 15; Length 352;
Best Local Similarity 95.0%; Pred. No. 2.1e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 KVNPHGMJLDEIANSKD 20
        ||||| ||||| ||||| |||||
DB      294 KVNPHGMJLDEIANSKD 313

RESULT 7
US-10-272-459-41
Sequence 41, Application US/10272459
Publication No. US20030124517A1
GENERAL INFORMATION:
APPLICANT: PICHANTES, Sergio
TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT
FILE REFERENCE: P017955.002 / 2301-17955
CURRENT APPLICATION NUMBER: US/10/272,459
CURRENT FILING DATE: 2002-10-15
NUMBER OF SEQ ID NOS: 50
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 41
LENGTH: 980
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41

Query Match          94.2%; Score 98; DB 15; Length 980;
Best Local Similarity 95.0%; Pred. No. 7e-08;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 KVNPHGMJLDEIANSKD 20
        ||||| ||||| ||||| |||||
DB      922 KVNPHGMJLDEIANSKD 941

RESULT 8
US-09-978-295A-616
Sequence 616, Application US/09978295A
Patent No. US2002015606A1
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Geo, Wei-Olang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
```

```

APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C11
CURRENT APPLICATION NUMBER: US/09/978,295A
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
```

;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080328  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080333  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080334  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/081070  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081049  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081071  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081195  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081955  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081819  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081952  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081838  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082566  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082569  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082704  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082804  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082700  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082797  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082796  
;; PRIOR FILING DATE: 1998-04-23  
;; PRIOR APPLICATION NUMBER: 60/083336  
;; PRIOR FILING DATE: 1998-04-27  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083392  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083495  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083496  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083499  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083554  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083558  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083559  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083500  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083742  
;; PRIOR FILING DATE: 1998-04-30  
;; PRIOR APPLICATION NUMBER: 60/084366  
;; PRIOR FILING DATE: 1998-05-05  
;; PRIOR APPLICATION NUMBER: 60/084414  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084441  
;; PRIOR FILING DATE: 1998-05-06

;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084639  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084640  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084598  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.3%; Score 43; DB 10; Length 98;  
Best Local Similarity 44.4%; Pred. No. 9.9;  
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 NFPHGMDLEIRANSKD 20  
Db 40 NLPFGVADLTQIDVNOVOD 57

RESULT 9  
US-09-978-697-616  
Sequence 616, Application US/09978697  
Patent No. US20020169284A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerlitsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James;  
APPLICANT: Paonli, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.

APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
Acids Encoding the Same  
FILE REFERENCE: P2630PIC27  
CURRENT APPLICATION NUMBER: US/09/978,697  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01

PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083316  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598

PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-5-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.3%; Score 43; DB 10; Length 98;  
Best Local Similarity 44.4%; Pred. No. 9.9;  
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 3 NPPHGMIDLEIRIAANSKD 20  
Db 40 NLEPGVADLTQIDVNOVD 57

RESULT 10  
US-09-978-192A-616  
Sequence 616, Application US/09978192A  
Patent No. US2002017753A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gertschen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C9

CURRENT APPLICATION NUMBER: US/09/978,192A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195



PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084643  
PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085582  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085700  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085689  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085580  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085573  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085704  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.3%; Score 43; DB 10; Length 98;  
Best Local Similarity 44.4%; Pred. No. 9.9;  
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 3 NPPHGLDDEEIRANSKD 20  
DB 40 NIPEGVADITQIDVAVQD 57

RESULT 11  
US-09-999-832a-616  
Sequence 616, Application US/09999832a  
Publication No. US20020192706A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Klavlin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James;  
APPLICANT: Paonli, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2650P1C63  
CURRENT FILING DATE: 2001-10-24  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249

PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15

1	PRIOR APPLICATION NUMBER: 60/081817
2	PRIOR FILING DATE: 1998-04-15
3	PRIOR APPLICATION NUMBER: 60/081819
4	PRIOR FILING DATE: 1998-04-15
5	PRIOR APPLICATION NUMBER: 60/081952
6	PRIOR FILING DATE: 1998-04-15
7	PRIOR APPLICATION NUMBER: 60/081838
8	PRIOR FILING DATE: 1998-04-15
9	PRIOR APPLICATION NUMBER: 60/082568
10	PRIOR FILING DATE: 1998-04-21
11	PRIOR APPLICATION NUMBER: 60/082569
12	PRIOR FILING DATE: 1998-04-21
13	PRIOR APPLICATION NUMBER: 60/082704
14	PRIOR FILING DATE: 1998-04-22
15	PRIOR APPLICATION NUMBER: 60/082804
16	PRIOR FILING DATE: 1998-04-22
17	PRIOR APPLICATION NUMBER: 60/082700
18	PRIOR FILING DATE: 1998-04-22
19	PRIOR APPLICATION NUMBER: 60/082797
20	PRIOR FILING DATE: 1998-04-22
21	PRIOR APPLICATION NUMBER: 60/082796
22	PRIOR FILING DATE: 1998-04-23
23	PRIOR APPLICATION NUMBER: 60/083336
24	PRIOR FILING DATE: 1998-04-27
25	PRIOR APPLICATION NUMBER: 60/083322
26	PRIOR FILING DATE: 1998-04-28
27	PRIOR APPLICATION NUMBER: 60/083392
28	PRIOR FILING DATE: 1998-04-29
29	PRIOR APPLICATION NUMBER: 60/083495
30	PRIOR FILING DATE: 1998-04-29
31	PRIOR APPLICATION NUMBER: 60/083496
32	PRIOR FILING DATE: 1998-04-29
33	PRIOR APPLICATION NUMBER: 60/083558
34	PRIOR FILING DATE: 1998-04-29
35	PRIOR APPLICATION NUMBER: 60/083539
36	PRIOR FILING DATE: 1998-04-29
37	PRIOR APPLICATION NUMBER: 60/083500
38	PRIOR FILING DATE: 1998-04-29
39	PRIOR APPLICATION NUMBER: 60/083742
40	PRIOR FILING DATE: 1998-04-30
41	PRIOR APPLICATION NUMBER: 60/084366
42	PRIOR FILING DATE: 1998-05-05
43	PRIOR APPLICATION NUMBER: 60/084414
44	PRIOR FILING DATE: 1998-05-06
45	PRIOR APPLICATION NUMBER: 60/084441
46	PRIOR FILING DATE: 1998-05-06
47	PRIOR APPLICATION NUMBER: 60/084637
48	PRIOR FILING DATE: 1998-05-07
49	PRIOR APPLICATION NUMBER: 60/084639
50	PRIOR FILING DATE: 1998-05-07
51	PRIOR APPLICATION NUMBER: 60/084640
52	PRIOR FILING DATE: 1998-05-07
53	PRIOR APPLICATION NUMBER: 60/084598
54	PRIOR FILING DATE: 1998-05-07
55	PRIOR APPLICATION NUMBER: 60/084600
56	PRIOR FILING DATE: 1998-05-07
57	PRIOR APPLICATION NUMBER: 60/084627
58	PRIOR FILING DATE: 1998-05-07
59	PRIOR APPLICATION NUMBER: 60/084643
60	PRIOR FILING DATE: 1998-05-07
61	PRIOR APPLICATION NUMBER: 60/085339
62	PRIOR FILING DATE: 1998-05-13
63	PRIOR APPLICATION NUMBER: 60/085338
64	PRIOR FILING DATE: 1998-05-13
65	PRIOR APPLICATION NUMBER: 60/085323
66	PRIOR FILING DATE: 1998-05-13
67	PRIOR APPLICATION NUMBER: 60/085582



[illegible]

? PRIOR APPLICATION NUMBER: 60/085580  
 ? PRIOR FILING DATE: 1998-05-15  
 ? PRIOR APPLICATION NUMBER: 60/085573  
 ? PRIOR FILING DATE: 1998-05-15  
 ? PRIOR APPLICATION NUMBER: 60/085704  
 ? PRIOR FILING DATE: 1998-05-15  
 ? PRIOR APPLICATION NUMBER: 60/085697

Query Match	41.3%	Score 43	DB 11	length 98
Best Local Similarity	44.4%	Pred. No.	9.9	
Matches	8	Conservative	3	Mismatches 7; Indels 0; Gaps 0;

```
QY      3 NRPHGMLDLEETIAANSKD 20
         | | | : | | : | | : |
Db      40 NLPEGVADLTQIDVNVQD 57
```

RESULT 13  
HE-00-078

US-09-9/8-608A-616  
; Sequence 616, Application US/09978608A  
; Publication No. US20030045462A1  
; Publication No. US20030045462A1

```

APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C22
CURRENT APPLICATION NUMBER: US/09/978,608A
CURRENT FILING DATE: 2001-10-16
NUMBER OF SEQ. ID NOS: 624
Prior Application removed - See file wrapper or Palm
SEQ ID NO 616
LENGTH: 98
TYPE: PRT
ORGANISM: Homo Sapien
US-09-978-608A-616

```

Query Match	41.3%	Score 43	DB 11	Length 98
Best Local Similarity	44.4%	Pred. No. 9.9		
Matches	8	Conservative	3	Mismatches 7
				Indels 0
				Gaps 0

```
QY      3 NFRHGMLDLEEIAANSKD 20
        | | | : | | | : |
Db      40 NLPEGVADLTQIDVNVQD 57
```

RESULT 14  
US-09-978-585A-616



;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080328  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080333  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/080334  
;; PRIOR FILING DATE: 1998-04-01  
;; PRIOR APPLICATION NUMBER: 60/081070  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081049  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081071  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081195  
;; PRIOR FILING DATE: 1998-04-08  
;; PRIOR APPLICATION NUMBER: 60/081203  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081229  
;; PRIOR FILING DATE: 1998-04-09  
;; PRIOR APPLICATION NUMBER: 60/081955  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081817  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081819  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081952  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/081838  
;; PRIOR FILING DATE: 1998-04-15  
;; PRIOR APPLICATION NUMBER: 60/082568  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082569  
;; PRIOR FILING DATE: 1998-04-21  
;; PRIOR APPLICATION NUMBER: 60/082704  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082804  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082700  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082797  
;; PRIOR FILING DATE: 1998-04-22  
;; PRIOR APPLICATION NUMBER: 60/082796  
;; PRIOR FILING DATE: 1998-04-23  
;; PRIOR APPLICATION NUMBER: 60/083336  
;; PRIOR FILING DATE: 1998-04-27  
;; PRIOR APPLICATION NUMBER: 60/083322  
;; PRIOR FILING DATE: 1998-04-28  
;; PRIOR APPLICATION NUMBER: 60/083392  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083495  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083496  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083499  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083554  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083558  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083559  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083500  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083742  
;; PRIOR FILING DATE: 1998-04-30  
;; PRIOR APPLICATION NUMBER: 60/084366  
;; PRIOR FILING DATE: 1998-05-05  
;; PRIOR APPLICATION NUMBER: 60/084414  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084441  
;; PRIOR FILING DATE: 1998-05-06

;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084639  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084640  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084598  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 41.38; Score 43; DB 11; Length 98;  
Best Local Similarity 44.48; Pred. No. 9.9;  
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
Oy 3 NPGHMDLEFIANSKD 20  
Db 40 NPEGVADLTQIDVYVD 57

Search completed: October 1, 2003, 10:37:52  
job time : 25.1176 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 ; Search time 43.5294 Seconds  
(without alignments)  
72.928 Million cell updates/sec

Title: US-09-171-432a-45  
Perfect score: 101  
Sequence: 1 DEERIANSKDFPNKSEFDL 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.\*
- 2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.\*
- 3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.\*
- 4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.\*
- 5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.\*
- 6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.\*
- 7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.\*
- 8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.\*
- 9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.\*
- 10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.\*
- 11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.\*
- 12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.\*
- 13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.\*
- 14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.\*
- 15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.\*
- 16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.\*
- 17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.\*
- 18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.\*
- 19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.\*
- 20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.\*
- 21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.\*
- 22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.\*
- 23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*
- 24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	20	18	AAW42928
2	101	100.0	21	22	ABG31729
3	101	100.0	22	7	ABP60066
4	101	100.0	22	11	AAK05697
5	101	100.0	22	18	AAW34074
6	101	100.0	22	21	AAI18607
7	101	100.0	22	21	AAI18608
8	101	100.0	22	21	AAI18609
9	101	100.0	22	23	ABG31727

10	101	100.0	22	23	ABG31728
11	101	100.0	22	23	ABG31729
12	101	100.0	22	23	ABE19899
13	101	100.0	22	24	ABU08639
14	101	100.0	22	24	ABU08640
15	101	100.0	22	24	ABU08641
16	53	52.5	20	18	AAW42927
17	53	52.5	21	22	ABG69444
18	46	45.5	42	24	ABB71556
19	44.5	44.1	17	22	ABB60536
20	44	43.6	80	23	ABP02822
21	43	42.6	152	18	AAW55492
22	43	42.6	157	18	AAW55327
23	43	42.6	350	20	AAV27183
24	43	42.6	536	21	AAW48090
25	43	42.6	1201	22	AAW90345
26	43	42.6	1201	22	ABW58421
27	42	41.6	89	22	AAU59799
28	42	41.6	177	21	AAW48507
29	42	41.6	226	21	AAW48506
30	42	41.6	265	21	AAW48505
31	42	41.6	365	18	AAW13493
32	42	41.6	365	20	AAV29461
33	42	41.6	402	22	ABW60260
34	42	41.6	425	22	ABW09060
35	42	41.6	439	24	ABP78252
36	42	41.6	445	22	AAW40763
37	42	41.6	445	22	AAW40764
38	42	41.6	445	22	AAW40765
39	42	41.6	538	21	AAV81713
40	42	41.6	538	24	ABU00469
41	42	41.6	641	22	AAW38978
42	42	41.6	659	23	AAO17073
43	42	41.6	678	22	AAW38977
44	42	41.6	698	22	AAW39032
45	42	41.6	705	22	AAW94135

## ALIGNMENTS

RESULT 1  
ID AAW42928 standard; peptide: 20 AA.  
XX  
AC AAW42928;  
XX  
DT 28-APR-1998 (first entry)  
XX  
DE Immunogenic Hepatitis A virus peptide YK-1328.  
XX  
DE Immunogenic Hepatitis A virus.  
XX  
KW Immune response; antibody.  
XX  
XX Synthetic.  
OS Hepatitis A virus.  
OS  
PN WO9740147-A1.  
XX  
PD 30-OCT-1997.  
XX  
PF 18-APR-1997; 97WO-US06891.  
XX  
PR 19-APR-1996; 96US-0015644.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Fields HA, Khudyakov YE.  
XX  
DR WPI; 1997-535831/49.  
XX  
PT Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an immune response to HAV in a mammal or to detect the presence of

antibodies against HAV in a mammal

Claim 18; Page 112; 140pp; English.

Peptides AAM42922-30 are immunogenic peptides corresponding to immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are substantially similar to a portion of the amino acid sequence of the P2A protein of HAV corresponding to amino acids 792-980. The present peptide is derived from amino acids 931-950, and has a reactivity of 12.5% with acute sera. Compositions containing the peptides can be used to induce an immune response to HAV in a mammal. The peptides can also be used to detect the presence of antibodies against HAV in mammalian serum. The administering the peptide to a mammal.

Sequence 20 AA:

Query Match 100.0%; Score 101; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.1e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEETIANSKDFPNMSETDL 20  
|||||  
1 DLEEIAANSKDFFPNMSETDL 20

RESULT 2  
AAB69445  
ID AAB69445 standard; Peptide; 21 AA.  
XX AAB69445;  
XX AC  
XX AD  
DT 20-APR-2001 (first entry)  
DE Synthetic HAV p2A peptide, SEQ ID NO: 45.  
XX  
XX Hepatitis A virus; HAV; immunogen; immunostimulant; vaccine; antigen; major structural capsid polypeptide; HAV antibody detection.  
KW  
XX  
OS Hepatitis A virus.  
OS Synthetic.  
XX  
XX WO200105824-A2.  
PD 25-JAN-2001.  
XX  
XX 14-JUL-2000; 2000WO-US19267.  
PF  
XX 15-JUL-1999; 99US-0144412.  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Fields HA, Khudjakov YE;  
PI  
DR WPI; 2001-112681/12.  
XX  
XX Claim 13; Page 97; 130pp; English.

The present sequence is one of a number of synthetic peptides which are immuno-reactive with hepatitis A virus (HAV) antibodies. The peptides comprise antigenic epitopes of the major structural capsid polypeptides or non-structural polypeptides of HAV with one or more glutamine molecules at the carboxy end of the peptide. The peptides are used to detect the presence of antibodies against HAV in mammalian serum, to detect the presence of HAV in a human or animal through the binding of the peptide to an antibody, to detect acute phase infection by detecting IgM antibodies in mammalian serum and detecting convalescence in a mammal. The peptides are used to detect or quantify HAV antibodies in samples in clinical or research-based assays using immunoblotting, fluorescence in situ hybridisation analysis, gel-mobility shift assays,

Query Match	Best Local Similarity	Score 101;	DB 22;	Length 21;
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 DLEEIAANSKDFPNMSEDTL 20			
Db	1 DLEEIAANSKDFPNMSEDTL 20			
RESULT 3				
AAP60066	AAP60066 standard; Protein; 2227 AA.			
XX	AAP60066;			
XX	25-MAR-2003 (updated)			
DT	26-JUN-1991 (first entry)			
XX	Sequence of viral I434 polypeptide encoded by the complete			
DE	nucleotide sequence of the HAV genome.			
XX	Diagnosis; vaccine; passive immunotherapy.			
KW	Hepatitis A virus.			
XX				
OS				
XX	Key	Location/Qualifiers		
FH	Region	1..245		
FT	Region	/label= P1.1A		
FT	Region	246..491		
FT	Region	/label= 1B		
FT	Region	492..836		
FT	Region	/label= 1C		
FT	Region	837..980		
FT	Region	/label= P2.2A		
FT	Region	981..1076		
FT	Region	/label= 2B		
FT	Region	1077..1422		
FT	Region	/label= 2C		
FT	Region	1423..1484		
FT	Region	/label= P3.3A		
FT	Region	1485..1507		
FT	Region	/label= 3B		
FT	Region	1508..1678		
FT	Region	/label= 3C		
FT	Region	1679..2227		
FT	Region	/label= 3D		
XX	EP199480-A.			
PN				
XX	29-OCT-1986.			
PD				
XX	03-APR-1986;	86EP-0302465.		
PF				
XX	03-APR-1985;	85US-0719329.		
PR				
XX	(CHIR ) CHIRON CORP.			
PA				
XX	Dina D, Potter SJ, Vannest GA, Caput D;			
PI				
XX	WPI; 1986-286213/44.			
DR	N-PSDB; AAN60080.			
XX				
XX	Hepatitis A virus nucleotide sequence and polypeptide - and use			
PT	in prodn. of vaccines and diagnostic probes			
XX				
XX	Claim 5; Fig 1; 18pp; English.			



Query Match	100.0%:	Score 101:	DB 7:	Length 2227:
Best Local Similarity	100.0%:	Pred. NO. 2.5e-07:		
Matches 20:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
OY	1 DLEEIAANSKDFPNMSEITDL 20			
DB	931 DLEEIAANSKDFPNMSEITDL 950			
SQLF 4				
ID	AA05697 standard; protein; 2227 AA.			
XX	AA05697;			
XX	25-MAR-2003 (updated)			
DT	15-AUG-1990 (first entry)			
XX	Attenuated hepatitis A virus.			
DE	Hepatitis A virus; vaccine; attenuated.			
KM	Hepatitis A virus, strain HM-175.			
XX				
OS				
XX				
XX	Key			
FT	Location/Qualifiers			
FT	1..23			
FT	/label-VP4 - 1A			
FT	24..245			
FT	/label-VP2 - 1B			
FT	246..491			
FT	/label-VP3 - 1C			
FT	492..791			
FT	/label-VP1 - 1D			
FT	792..980			
FT	/label-2A			
FT	981..1087			
FT	/label-2B			
FT	1088..1422			
FT	/label-2C			
FT	1423..1496			
FT	/label-3A			
FT	1497..1519			
FT	/label-3B - VPg			
FT	1520..1738			
FT	/label-3C			
FT	1739..2227			
FT	/label-3D			
XX				
PN	US4894228-A.			
XX				
PD	16-JAN-1990.			
XX				
PE	12-JUL-1988; 88US-0217824.			
XX				
PR	12-JUL-1988; 88US-0217824.			
XX				
XX	12-JUL-1988; 88US-0652967.			
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICE.			
PI	Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstone SM;			
TI	Dammer KU, Gust ID;			

```

xx  WP1: 1990-075557/10.
DR  N-PSDB: AA003512.
xx
xx  Vaccine against hepatitis A virus infection - comprises novel
PT  attenuated hepatitis A virus strain.
xx
xx  Claim 1; Fig 1; 18pp; English.
xx
xx  The attenuated HAV is useful for inducing protective immunity against
CC  HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by
CC  several nucleotide changes distributed throughout the genome, is
CC  attenuated for chimpanzees, elicits serum neutralising antibodies, and is
CC  suitable for use as an HAV vaccine. It is noted that not all the changes
CC  are necessary for attenuation and use as a vaccine.
CC  (Updated on 25-MAR-2003 to correct PA field.)
CC  (Updated on 25-MAR-2003 to correct PI field.)
xx
xx  Sequence 2227 AA;
SO
xx
xx  Query Match 100.0%; Score 101; DB 11; Length 2227;
xx  Best Local Similarity 100.0%; Pred. No. 2.5e-07;
xx  Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
xx
xx  1 DLEETAAASKDPNNSETDL 20
xx  |||
xx  931 DLEETAAASKDPNNSETDL 950
xx
xx  RESULT 5
xx  AAWM34074
xx  AAWM34074 standard; Protein; 2227 AA.
xx
xx  AAW34074;
xx
xx  27-APR-1998 (first entry)
xx
xx  Hepatitis A virus HM-175 protein sequence.
xx
xx  HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;
xx  infection; vaccine.
xx
xx  OS
xx  Hepatitis A virus HM-175.
xx
xx  Key
xx  FH Protein
xx  FT /label- VP4
xx  FT 24..245
xx  FT /label- VP2
xx  FT 246..491
xx  FT /label- VP3
xx  FT 492..791
xx  FT /label- VP1
xx  FT 792..980
xx  FT /label- 2A
xx  FT 981..1087
xx  FT /label- 2B
xx  FT 1088..1422
xx  FT /label- 2C
xx  FT 1423..1496
xx  FT /label- 3A
xx  FT 1497..1519
xx  FT /label- 3B
xx  FT 1520..1738
xx  FT /label- 3C
xx  FT 1739..2227
xx  FT /label- 3D
xx
xx  MO9740166-A2.
xx
xx  30-OCT-1997.
xx
xx  18-APR-1997; 97WO-US06506.
xx

```

```

XX 19-APR-1996; 96US-0015642.
XX (USSH ) US SEC DEPT HEALTH.
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX Emerson SU, Purcell RH, Raychaudhuri G;
XX WPI; 1997-535850/49.
XX N-PSDB; AAT93023.
XX Human attenuated HAV genome containing simian HAV 2C gene - useful
XX as vaccines against HAV infection
XX Disclosure: Fig 13A-D; 66pp; English.
XX
XX This protein sequence is encoded by the human hepatitis A virus
XX (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain
XX HAV/7 is obtained by passage of HM-175 in African Green Monkey
XX kidney cells. A claimed DNA construct (1) comprises a genome of
XX HAV, where the genome is a human attenuated HAV genome in which a
XX region of the 2C gene has been replaced by a corresponding region
XX from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The
XX region of the 2C gene from AGM-27 contained in the construct
XX preferably encodes amino acids 120-328 of the 2C protein, amino
XX acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA
XX transcript of (1); (2) a cell transfected with (1) or the RNA
XX transcript of (1); (3) a HAV genome as above; (4) antibodies to the
XX HAV of (3); and (5) a host cell containing the HAV of (3). (1) or
XX its RNA transcript, can be used as a vaccine for preventing HAV in
XX a mammal. (1) or the RNA transcript can also be used to stimulate
XX the production of protective antibodies in the mammal.
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 101; DB 18; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2,5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DLEFIANSKDFPMSETDL 20
DB 931 DLEFIANSKDFPMSETDL 950

```

```

DR N-PSDB; AAA75476.
XX
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell
XX line useful as vaccine for protecting humans against hepatitis A virus
XX infection, has modified genome compared to wild type
XX
XX Disclosure: Fig 6A-K; 72pp; English.
XX
XX The present sequence is derived from a wild type hepatitis A virus
XX (HAV) strain HM-174. The sequence is modified to produce HAV which
XX are adapted to growth in the human fibroblast-like cell line MRC-5.
XX The HAV is able to propagate in MRC-5 cells and retain appropriate
XX attenuation. It is useful as a live vaccine for prophylaxis of
XX hepatitis A in humans and other primates.
XX
XX Sequence 2227 AA;
SQ
Query Match 100.0%; Score 101; DB 21; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2,5e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DLEFIANSKDFPMSETDL 20
DB 931 DLEFIANSKDFPMSETDL 950

```

Query Match	100.0%	Score 101	DB 21	Length 2227
Best Local Similarity	100.0%	Pred. No. 2.5e-07		
Matches	20	Conservative 0	Mismatches 0	Indels 0
Qy	1	DLEEAANSKDPNNSETDL	20	
Db	931	DLEEAANSKDPNNSETDL	950	

RESULT 8  
AAB18609  
ID AAB18609 standard; Protein; 2227 AA

DT	15-JAN-2001 (first entry)
XX	
DE	Amino acid sequence of live attenuated Hepatitis A virus 4380.
NAV	strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection; HAV 4380.

XX Hepatitis A virus.  
OS US6113912-A.  
XX  
PN 05-SEP-2000.  
XX  
PD 07-JUN-1995; 95US-0475886.  
XX  
PF 18-SEP-1992; 92US-0947338.  
XX 17-SEP-1993; 93WO-US08610.  
PR 10-MAR-1995; 95US-0397232.  
PR  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES

PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW.  
XX  
DR WPI: 2000-586464/55.  
DR N-PSDB; AAA75478.

PT Novel live hepatitis A virus adapted to growth in human fibroblast cell  
PT line useful as vaccine for protecting humans against hepatitis A virus  
PT infection, has modified genome compared to wild type -  
xx  
PS Disclosure; Columns 93-104; 72pp: English.  
xx

The present sequence is derived from a live attenuated hepatitis A virus (HAV) of the invention, designated HAV 4380. The sequence is produced by modifying wild type HAV strain HM-174. The HAV of the invention are adapted to growth in the human fibroblast-like cell line MRC-5. The HAV is able to propagate in MRC-5 cells and retain appropriate attenuation. It is useful as a live vaccine for prophylaxis of hepatitis A in humans and other primates.

**Sequence 2227 AA;**

Query Match	100.0%	Score 101	DB 21	Length 2227
Best Local Similarity	100.0%	Pred. NO.	2.5e-07	
Matches 20	Conservative 0	Mismatches 0	Indels 0	Gaps 0

```

QY      1 DLEETIAANSKDFPPNMSSETDL 20
        |||||
DB      931 DLEETIAANSKDFPPNMSSETDL 950

```

RESULT 9	
ABG31727	
ID	ABG31727 standard; Protein; 2227 AA
XX	
AC	ABG31727;
XX	

DT	29-NOV-2002 (first entry)
XX	
XX	Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.
DE	
XX	Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.
KW	
XX	Hepatitis A virus strain HM-175.
DS	

23-JUL-2002.

PF 31-AUG-2000; 2000US-0653499.

PR 07-JUN-1995; 95US-0475886.

PR 17-SEP-1993; 93WO-US08610.

PA (USSH ) US DEPT HEALTH &amp; HUMAN SERVICES.

XX

XX

DR N-PSDB; ABS52787.

PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth

XX XX

XX

CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cells).

CC hepatitis A virus infection. This sequence represents a hepatitis A virus

XX

**SQ Sequence 2227 AA;**

Query Match	100.08;	Score 101;	DB 23;	Length 2227;
-------------	---------	------------	--------	--------------

Matches	20;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	----	--------	----	------	----

```
QY      1 DLEEIAANSKDFPNMSETDL 20  
        |||||  
Db     931 DLEEIAANSKDFPNMSETDL 956
```

RESULT 10  
ABG31728

ID ABG31728 standard; Protein; 2227 AA.

AC ABG31728

DT 29-NOV-2002 (first entry)

Hepatitis A virus mutant strain HM-175/7 (pHAV/7) polypeptide.

KW Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;

2000

OS Synthetic.

FH	Key	Location/Qualifiers
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

	/label= Wild-type Lys substituted by Arg
FT	
FT	

FT	/note=	*Wild-type Glu substituted by Val
FT	2166	001

ET	note=	*Wild-type Asn substituted by Ser
1053		
1054		
1055		
1056		
1057		
1058		
1059		
1060		
1061		
1062		
1063		
1064		
1065		
1066		
1067		
1068		
1069		
1070		
1071		
1072		
1073		
1074		
1075		
1076		
1077		
1078		
1079		
1080		
1081		
1082		
1083		
1084		
1085		
1086		
1087		
1088		
1089		
1090		
1091		
1092		
1093		
1094		
1095		
1096		
1097		
1098		
1099		
1100		
1101		
1102		
1103		
1104		
1105		
1106		
1107		
1108		
1109		
1110		
1111		
1112		
1113		
1114		
1115		
1116		
1117		
1118		
1119		
1120		
1121		
1122		
1123		
1124		
1125		
1126		
1127		
1128		
1129		
1130		
1131		
1132		
1133		
1134		
1135		
1136		
1137		
1138		
1139		
1140		
1141		
1142		
1143		
1144		
1145		
1146		
1147		
1148		
1149		
1150		
1151		
1152		
1153		
1154		
1155		
1156		
1157		
1158		
1159		
1160		
1161		
1162		
1163		
1164		
1165		
1166		
1167		
1168		
1169		
1170		
1171		
1172		
1173		
1174		
1175		
1176		
1177		
1178		
1179		
1180		
1181		
1182		
1183		
1184		
1185		
1186		
1187		
1188		
1189		
1190		
1191		
1192		
1193		
1194		
1195		
1196		
1197		
1198		
1199		
1200		

ET /note= "Wild-type Ala substituted by Val

```

FT Misc-difference 1062 /note- "Wild-type Gly substituted by Ala"
FT Misc-difference 1118 /note- "Wild-type Lys substituted by Met"
FT Misc-difference 1151 /note- "Wild-type Glu substituted by Lys"
FT Misc-difference 1163 /note- "Wild-type Phe substituted by Ser"
FT Misc-difference 1277 /note- "Wild-type Val substituted by Ile"
FT Misc-difference 1500 /note- "Wild-type His substituted by Tyr"
FT Misc-difference 1805 /note- "Wild-type Asp substituted by Asn"
FT Misc-difference 1930 /note- "Wild-type Ser substituted by Thr"
FT
FT
FT US6423318-B1.
FT
FT
FT 23-JUL-2002.
FT
FT 31-AUG-2000; 2000US-0653499.
FT
FT 07-JUN-1995; 95US-0475886.
FT 17-SEP-1993; 93US-0397232.
FT 17-SEP-1993; 93WO-US08610.
FT
FT (USSH ) US DEPT HEALTH & HUMAN SERVICES.
FT (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
FT
FT Funckouser AM, Emerson SU, Purcell RH, D'Hondt E;
FT
FT WPI; 2002-680946/73.
FT N-PSDB; ABSS2788.
FT
FT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
FT in MRC-5 cells, useful for preparing a vaccine against HAV infection -
FT
FT
FT Example 3; Column 67-78; 71pp; English.
FT
FT The invention relates to a polynucleotide which encodes a hepatitis A
FT virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
FT line). The polynucleotide is useful for preparing a vaccine against
FT hepatitis A virus infection. This sequence represents a hepatitis A virus
FT CC mutant strain HM-175/7 (pHAV/7) polypeptide.
FT CC
FT
FT Sequence 2227 AA;
FT
FT Query Match 100.0%; Score 101; DB 23; Length 2227;
FT Best Local Similarity 100.0%; Pred. No. 2.5e-07;
FT Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
FT
FT 1 DLEFIANSKDFPMSETDL 20
FT
FT 931 DLEFIANSKDFPMSETDL 950
FT
FT
FT RESULT 11
FT ABG31729
FT ID ABG31729 standard; Protein; 2227 AA.
FT
FT AC ABG31729;
FT
FT 29-NOV-2002 (first entry)
FT
FT Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
FT
FT Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
FT KM HAV 4380.
FT
FT Hepatitis A virus strain HM-175.
FT OS
FT US6423318-B1.
FT PN

```

```

XX 23-JUL-2002.
XX
XX 31-AUG-2000; 2000US-0653499.
XX
XX 07-JUN-1995; 95US-0475886.
XX 17-SEP-1993; 93US-0397232.
XX 17-SEP-1993; 93WO-US08610.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Funckouser AM, Emerson SU, Purcell RH, D'Hondt E;
XX
XX WPI; 2002-680946/73.
XX N-PSDB; ABSS2789.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX in MRC-5 cells, useful for preparing a vaccine against HAV infection -
XX
XX
XX Disclosure; Column 93-104; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
XX virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX line). The polynucleotide is useful for preparing a vaccine against
XX hepatitis A virus infection. This sequence represents an attenuated
XX hepatitis A virus 4830 polypeptide.
XX
XX Sequence 2227 AA;
XX
XX Query Match 100.0%; Score 101; DB 23; Length 2227;
XX Best Local Similarity 100.0%; Pred. No. 2.5e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 DLEFIANSKDFPMSETDL 20
XX
XX 931 DLEFIANSKDFPMSETDL 950
XX
XX
XX RESULT 12
XX AAE19899
XX ID AAE19899 standard; Protein; 2227 AA.
XX
XX AC AAE19899;
XX
XX 18-JUN-2002 (first entry)
XX
XX Hepatitis A virus (HAV) protein.
XX
XX Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
XX KM cytosstatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
XX Hepatitis A virus.
XX OS
XX WO200213855-A2.
XX
XX 21-FEB-2002.
XX
XX 15-AUG-2001; 2001WO-IB01808.
XX
XX 17-AUG-2000; 2000US-225767P.
XX 29-AUG-2000; 2000US-228175P.
XX 03-NOV-2000; 2000US-0705547.
XX
XX (TRIP-) TRIPEP AB.
XX
XX Saliberg M, Hultgren C;
XX
XX WPI; 2002-241837/29.
XX N-PSDB; AAD31766.
XX
XX Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX

```

PT epitope present in hepatitis C virus -

XX Claim 11; Page 82-87; 120pp; English.

XX The invention relates to a composition comprising ribavirin and an

CC antigen preferably non structural 3 protein (NS3)/4A fragment of

CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV

CC sequence. The composition is useful for enhancing an immune response to

CC a hepatitis C antigen in humans, domestic, sport or pet species and as

CC vaccines for treating and preventing HCV infections. The composition is

CC also useful for treating viral, bacterial, fungal diseases and cancer.

CC The present sequence is hepatitis A virus (HAV) protein.

XX

SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 23; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.5e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEETIANSKDFPNMSETDL 20

|||||

931 DLEETIANSKDFPNMSETDL 950

RESULT 13

ABU08639

ID ABU08639 standard; Protein; 2227 AA.

XX

AC ABU08639;

XX

DT 03-JUN-2003 (first entry)

XX

DE Wild type human hepatitis A virus strain HM-175.

XX

KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;

KM vaccine; MRC-5 cell; hepatitis infection.

XX

OS Hepatitis A virus strain HM-175.

XX

PN US2002176869-A1.

XX

PD 28-NOV-2002.

XX

PF 29-APR-2002; 2002US-0135988.

XX

PR 07-JUN-1995; 95US-0475886.

PR 31-AUG-2000; 2000US-0653499.

DR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-0508610.

PR 17-APR-1995; 95US-0397232.

XX

PA (FUNK/) FUNKHOUSER A W.

PA (EMER/) EMERSON S U.

PA (PURC/) PURCELL R H.

PA (DHON/) D'HONDT E.

XX

PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;

XX

DR N-PSDB; ABX93473.

XX

PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,

PT and disease -

XX

PS Disclosure: Fig 6; 70pp; English.

XX

CC The invention describes a live hepatitis A virus (HAV) adapted to growth

CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as

CC a vaccine for protecting primates against hepatitis infection and

CC disease. This is the amino acid sequence of wild type human hepatitis A

CC virus strain HM-175.

XX

SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 24; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.5e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20

|||||

931 DLEETIANSKDFPNMSETDL 950

RESULT 14

ABU08640

ID ABU08640 standard; Protein; 2227 AA.

XX

AC ABU08640;

XX

DT 03-JUN-2003 (first entry)

XX

DE Attenuated (pass35) hepatitis A virus strain HM-175.

XX

KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;

KM vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.

XX

OS Hepatitis A virus strain HM-175.

XX

PN US2002176869-A1.

XX

PD 28-NOV-2002.

XX

PF 29-APR-2002; 2002US-0135988.

XX

PR 07-JUN-1995; 95US-0475886.

PR 31-AUG-2000; 2000US-0653499.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-0508610.

PR 17-APR-1995; 95US-0397232.

XX

PA (FUNK/) FUNKHOUSER A W.

PA (EMER/) EMERSON S U.

PA (PURC/) PURCELL R H.

PA (DHON/) D'HONDT E.

XX

PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;

XX

DR N-PSDB; ABX93474.

XX

PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,

PT and disease -

XX

PS Example 3; Fig 6; 70pp; English.

XX

CC The invention describes a live hepatitis A virus (HAV) adapted to growth

CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as

CC a vaccine for protecting primates against hepatitis infection and

CC disease. This is the amino acid sequence of an attenuated (pass 35)

CC human hepatitis A virus strain HM-175.

XX

SQ Sequence 2227 AA;

Query Match 100.0%; Score 101; DB 24; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.5e-07;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETIANSKDFPNMSETDL 20

|||||

931 DLEETIANSKDFPNMSETDL 950

RESULT 15

ABU08641

ID AB008641 standard; Protein; 2227 AA.  
 AC AB008641;  
 XX  
 XX  
 DT 03-JUN-2003 (first entry)  
 XX  
 DE Attenuated hepatitis A virus (4380) strain HM-175.  
 XX  
 KM Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;  
 KM vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.  
 XX  
 OS Hepatitis A virus strain HM-175.  
 XX  
 PN US2002176869-A1.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PE 29-APR-2002; 2002US-0135988.  
 XX  
 PF 07-JUN-1995; 95US-0475886.  
 31-AUG-2000; 2000US-0653499.  
 18-SEP-1992; 92US-0947338.  
 17-SEP-1993; 93WO-US08610.  
 PR 17-APR-1995; 95US-0397232.  
 XX  
 PA (FUNK/) FUNKHOUSER A W.  
 PA (EMER/) EMERSON S U.  
 PA (PURC/) PURCELL R H.  
 PA (DHON/) D'HONDT E.  
 XX  
 PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;  
 XX  
 DR WPI; 2003-352605/33.  
 DR N-PSDB; ABX93475.  
 XX  
 PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,  
 PT useful in vaccines for protecting primates against hepatitis infection  
 PT and disease  
 XX  
 PS Disclosure; Page 45-51; 70pp; English.  
 XX  
 CC The invention describes a live hepatitis A virus (HAV) adapted to growth  
 CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as  
 CC a vaccine for protecting primates against hepatitis infection and  
 CC disease. This is the amino acid sequence of an attenuated human  
 CC hepatitis A virus (4380) strain HM-175.  
 CC  
 SO Sequence 2227 AA;  
 XX  
 XX Query Match 100.0%; Score 101; DB 24; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-07;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 DLEBIANSKDFPMSETDL 20  
 DB 931 DLEBIANSKDFPMSETDL 950

Search completed: October 1, 2003, 09:56:44  
 Job time : 44.5294 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 12.8235 Seconds  
(without alignments)  
149.988 Million cell updates/sec

Title: US-09-171-432a-45

Perfect score: 101  
Sequence: 1 DLEEIANSKDFPNMSETDL 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	101	100.0	2227 1	GNMYHM genome polyprotein
2	101	100.0	2227 1	GNMYHM genome polyprotein
3	101	100.0	2227 1	GNMYHM genome polyprotein
4	101	100.0	2227 1	GNMYHM genome polyprotein
5	98	97.0	2230 1	GNMYHM genome polyprotein
6	50	49.5	736 2	D90574 genome polyprotein
7	49	48.5	400 2	F88931 hypothetical prote
8	45.5	45.0	930 2	A84668 Argonaute (AGO)-1
9	45.5	45.0	1621 2	T15264 hypothetical prote
10	44	43.6	740 1	FOLJHD gag polyprotein -
11	44	43.6	927 2	T38127 phosphoprotein - f
12	44	43.6	1046 2	S67786 hypothetical prote
13	43.5	43.1	754 1	BABOH peptide-aspartate
14	43.5	43.1	1451 2	S65571 pattern formation
15	43.5	43.1	5107 2	T29144 partial CDS - Caen
16	43	42.6	152 2	A64610 hypothetical prote
17	43	42.6	152 2	A71904 hypothetical prote
18	43	42.6	231 2	B70407 rhodanese family p
19	43	42.6	285 2	B87426 rhodanese family p
20	43	42.6	454 2	T02100 hypothetical prote
21	43	42.6	598 2	H71336 probable cell divi
22	43	42.6	1072 2	A84112 alkaline amylopull
23	43	42.6	1201 2	T08603 kinesin-related pr
24	42.5	42.1	913 2	T15278 hypothetical prote
25	42.5	42.1	4717 2	T41581 hypothetical colle
26	42	41.6	289 2	T23342 hypothetical prote
27	42	41.6	365 2	C81050 cytochrome c oxida
28	42	41.6	365 2	F81826 probable cytochrom
29	42	41.6	538 2	G95015 ABC transporter, A

30	42	41.6	538 2	A97889 hypothetical prote
31	42	41.6	1112 2	T47784 hypothetical prote
32	42	41.6	1946 2	AE1449 hypothetical prote
33	41.5	41.1	236 2	A81283 probable periplasm
34	41.5	41.1	694 2	S41868 DNA-directed RNA p
35	41.5	41.1	1378 2	A81393 hypothetical prote
36	41	40.6	94 2	D70245 hypothetical prote
37	41	40.6	108 2	AF2042 hypothetical prote
38	41	40.6	133 2	A29174 hypothetical prote
39	41	40.6	139 2	T26224 hypothetical prote
40	41	40.6	230 2	T45754 hypothetical prote
41	41	40.6	262 2	E89760 conserved hypotet
42	41	40.6	292 2	F86431 hypothetical prote
43	41	40.6	295 2	H83642 probable 2-OH-laur
44	41	40.6	346 2	G64182 lipopolysaccharide
45	41	40.6	368 2	G97291 hypothetical prote

## ALIGNMENTS

## RESULT 1

GNMYHM genome polyprotein - human hepatitis A virus (strain HM-175, wild type)  
N: Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core  
B: RNA-directed RNA polymerase (EC 2.7.7.48), protein 3D  
C: Species: human hepatitis A virus  
A: Note: host Homo sapiens (man)  
C: Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 16-Jul-1999  
C: Accession: A25981  
J: Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.  
J: Virol. 61, 50-59, 1987  
A: Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with  
A: Reference number: A25981; MID:87061253; PMID:3023706  
A: Accession: A25981  
A: Molecule type: genomic RNA  
A: Residues: 1-2227 <CON>  
C: Cross-references: EMBL:M14707; NID:G329582; PIDN:AAA5465.1; PID:G329583  
C: Superfamily: hepatitis A virus genome polyprotein  
C: Keywords: coat protein; core protein; cytosine proteinase; hydrolase; nucleotidyltr  
F: 1-23/Product: coat protein 1A #status predicted <VP1>  
F: 24-245/Product: coat protein 1B #status predicted <VP2>  
F: 246-491/Product: coat protein 1C #status predicted <VP3>  
F: 492-791/Product: coat protein 1D #status predicted <VP4>  
F: 792-980/Product: core protein 2A #status predicted <C2A>  
F: 981-1087/Product: core protein 2B #status predicted <C2B>  
F: 1088-1432/Product: core protein 2C #status predicted <C2C>  
F: 1423-1486/Product: protein 3A #status predicted <C3A>  
F: 1497-1519/Product: protein 3B #status predicted <C3B>  
F: 1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>  
F: 1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 101; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLEEIANSKDFPNMSETDL 20  
Db 931 DLEEIANSKDFPNMSETDL 950

RESULT 2  
GNMYHM genome polyprotein - human hepatitis A virus  
N: Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core  
NA polymerase (EC 2.7.7.48), protein 3D  
C: Species: human hepatitis A virus  
A: Note: host Homo sapiens (man)  
C: Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Jul-1999  
C: Accession: A03903  
R: Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N  
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985  
A: Title: Primary structure and gene organization of human hepatitis A virus.





F:28-249/Product: coat protein 1B #status predicted <C1B>  
 F:250-495/Product: coat protein 1C #status predicted <C1C>  
 F:496-795/Product: coat protein 1D #status predicted <C1D>  
 F:796-984/Product: core protein 2A #status predicted <C2A>  
 F:985-1091/Product: core protein 2B #status predicted <C2B>  
 F:1092-1426/Product: core protein 2C #status predicted <C2C>  
 F:1427-1498/Product: protein 3A #status predicted <P3A>  
 F:1499-1521/Product: protein 3B #status predicted <P3B>  
 F:1522-1741/Product: protein 3C #status predicted <P3C>  
 F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 97.0%; Score 98; DB 1; Length 2230;  
 Best Local Similarity 95.0%; Pred. No. 6e-07;  
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEIANSKDFPNNSETDL 20  
 |||||:|||||:|||||:  
 Db 935 DLEIANSKDFPNNSETDL 954

ULT 6  
 0574

hypothetical protein MYPV\_5000 [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
 C:Species: Mycoplasma pulmonis  
 C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 03-Aug-2001  
 C:Accession: D90574  
 R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Moszer, I.;  
 Nucleic Acids Res. 29, 2145-2153, 2001  
 A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm  
 A:Reference number: A95512; MUID:21267165; PMID:11353084  
 A:Accession: D90574  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-736 <KUR>  
 A:Cross-references: GB:AL445566; PID:g14089914; PIDN:CAC13673.1; GSPDB:GN00153  
 A:Experimental source: strain UAB CTIP  
 C:Genetics:  
 A:Gene: MYPV\_5000  
 A:Genetic code: SGC3

Query Match 49.5%; Score 50; DB 2; Length 736;  
 Best Local Similarity 55.0%; Pred. No. 8.3;  
 Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 DLEIANSKDFPNNSETDL 20  
 |||||:|||||:|||||:  
 Db 218 DLYSIYNEKDAPEISEDL 237

SULF 7  
 F88931

protein RL1G11.1 [imported] - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
 C:Accession: F88931  
 R:Anonymous, The C. elegans Sequencing Consortium.  
 Science 282, 2012-2018, 1998  
 A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog  
 A:Reference number: A75000; MUID:99069613; PMID:9851916  
 A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
 A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
 A:Accession: F88931  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-400 <STO>  
 A:Cross-references: GB:chr\_V; PIDN:AAC69076.1; PID:g2384851; GSPDB:GN00023; CESP:RL1G11.  
 C:Genetics:  
 A:Gene: RL1G11.1  
 A:Map position: 5

Query Match 48.5%; Score 49; DB 2; Length 400;  
 Best Local Similarity 47.4%; Pred. No. 5.9;

Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
 OY 2 LBEIANSKDFPNNSETDL 20  
 |||||:|||||:|||||:  
 Db 164 LBEIANSKDFPNNSETDL 182

RESULT 8

Argonate (AGOL)-like protein [imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
 C:Accession: A84668  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.  
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,  
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
 Nature 402, 761-768, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487; PMID:10617197  
 A:Accession: A84668  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-930 <STO>

A:Cross-references: GB:AE002093; NID:g3885334; PIDN:AAC77862.1; GSPDB:GN00139

C:Superfamily: rabbit translation initiation factor eIF-2C  
 A:Gene: At2g27040  
 A:Map position: 2

Query Match 45.0%; Score 45.5; DB 2; Length 930;  
 Best Local Similarity 42.3%; Pred. No. 57;  
 Matches 11; Conservative 4; Mismatches 2; Indels 9; Gaps 1;

OY 2 LBEIANSKDFPNNSETDL 18  
 |||||:|||||:|||||:  
 Db 153 LBEIANSKDFPNNSETDL 178

RESULT 9

hypothetical protein F59E12.9 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T15264  
 R:Johnson, D.  
 submitted to the EMBL Data Library, May 1997  
 A:Description: The sequence of C. elegans cosmid F59E12.  
 A:Reference number: Z18318  
 A:Accession: T15264  
 A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-1621 <JOH>  
 A:Cross-references: EMBL:AF003366; NID:g2088833; PID:g2088843; PIDN:AAB54259.1; GSPDB  
 A:Experimental source: strain Bristol N2; clone F59E12  
 C:Genetics:  
 A:Gene: CESP:F59E12.9  
 A:Map position: 2

A:Introns: 30/3; 55/1; 200/2; 299/2; 327/2; 369/3; 589/3; 860/1; 1278/1; 1547/

Query Match 45.0%; Score 45.5; DB 2; Length 1621;  
 Best Local Similarity 42.9%; Pred. No. 1.1e+02;  
 Matches 9; Conservative 8; Mismatches 1; Indels 3; Gaps 1;

OY 2 LBEIANSKDFPNNSETDL 19  
 |||||:|||||:|||||:  
 Db 985 LBEIANSKDFPNNSETDL 1005

RESULT 10

gag polyprotein - squirrel monkey retrovirus SMRV-H  
 FOLJHD  
 N:Contains: core protein p16; core protein p19; probable core protein p10; probable c  
 C:Species: squirrel monkey retrovirus SMRV-H

C>Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Jul-1999  
C/Accession: A31827  
R/Oda, T.; Ikeda, S.; Matanabe, S.; Hatsushika, M.; Akiyama, K.; Mitsunobu, F.  
Virology 167, 468-476, 1988  
A/Title: Molecular cloning, complete nucleotide sequence, and gene structure of the pro  
A/Reference number: A31827; MUID:9073750; PMID:3201749  
A/Accession: A31827  
A/Molecule type: DNA  
A/Residues: 1-740 <ODA>  
A/Cross-references: GB:M23385; NID:g332626; PIDN:AAA6451.1; PID:g807672  
C/Genetics:  
A/Genes: gag  
C/Superfamily: AIDS-related virus gag polyprotein  
C/Keywords: core protein, polyprotein  
F:1163/Product: core protein p19 #status predicted <CP9>  
F:114-318/Product: core protein p16 #status predicted <CP6>  
F:319-648/Product: core protein p35 #status predicted <CP5>  
F:649-740/Product: core protein p10 #status predicted <CP1>

Query Match 43.6%; Score 44; DB 1; Length 740;  
Best Local Similarity 50.0%; Pred. No. 76;  
Matches 10; Conservative 4; Mismatches 4; Indels 2; Gaps 1;  
OY 1 DLEBIAA-NSKDFPNMSET 18  
Db 214 DLEBAAQYNNPDPQTLNT 233

## RESULT 11

T38127  
Phosphoprotein - fission yeast (Schizosaccharomyces pombe)  
C/Species: Schizosaccharomyces pombe  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 02-Aug-2002  
C/Accession: T38127; A57087; S54119  
R/Badcock, K.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, May 1997  
A/Reference number: Z21772  
A/Accession: T38127  
A/Status: preliminary; translated from GB/EMBL/DBDJB  
A/Molecule type: DNA  
A/Residues: 1-927 <BAD>  
A/Cross-references: EMBL:Z95334; PIDN:CAB08599.2; GSPDB:GN00066; SPDB:SPAC2068.05C  
R/Finkhauser, C.; Raymond, A.; Cerutti, L.; Utzig, S.; Hofmann, K.; Simanis, V.  
Cell 82, 435-444, 1995  
A/Title: The Schizosaccharomyces pombe cdc15 gene is a key element in the reorganization  
A/Reference number: A57087; MUID:95360987; PMID:7634333  
A/Accession: A57087  
A/Status: preliminary; nucleic acid sequence not shown  
A/Molecule type: DNA  
A/Residues: MTRSLQ', 28-927 <FAN>  
A/Cross-references: GB:X86179  
R/Finkhauser, C.; Raymond, A.; Cerutti, L.; Simanis, V.  
submitted to the EMBL Data Library, April 1995  
A/Description: The cdc15 gene is a key element in F-actin reorganisation at mitosis.  
A/Reference number: S54119  
A/Accession: S54119  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 'MTRSLQ', 28-870, 'GKFIKEL' <FAZ>  
A/Cross-references: EMBL:X86179  
C/Genetics:  
A/Genes: SPDB:SPAC2068.05c; cdc15  
A/Map position: 1  
A/Intons: 27/3; 58/2; 871/1  
C/Superfamily: fission yeast scd2 protein; SH3 homology  
C/Keywords: mitosis; phosphoprotein  
F:873-924/Domain: SH3 homology <SH3>

Query Match 43.6%; Score 44; DB 2; Length 927;  
Best Local Similarity 47.1%; Pred. No. 98;  
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 2 LEEIANSKDFPNMSET 18  
Db 69 LOELIASSADIPVGST 85

## RESULT 12

S67786  
hypothetical protein YDL223c - yeast (Saccharomyces cerevisiae)  
N/Alternate names: hypothetical protein D0843  
C/Species: Saccharomyces cerevisiae  
C/Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 19-Apr-2002  
C/Accession: S67786  
R/Rasmussen, S.W.  
submitted to the Protein Sequence Database, July 1996  
A/Reference number: S67786  
A/Accession: S67786  
A/Molecule type: DNA  
A/Residues: 1-1046 <RAS>  
A/Cross-references: EMBL:Z74271; NID:g1431374; PID:e253363; PID:g1431375; GSPDB:GN000  
A/Experimental source: strain S288C  
C/Genetics:  
A/Genes: MIPS:YDL223c  
A/Cross-references: SGD:S0002382  
A/Map position: 4L

Query Match 43.6%; Score 44; DB 2; Length 1046;  
Best Local Similarity 50.0%; Pred. No. 1,1e+02;  
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 3 EELIANSKDFPNMSET 18  
Db 733 QDIASDAKDFNNPET 748

## RESULT 13

BAB0H  
peptide-aspartate beta-dioxygenase (EC 1.14.11.16) - bovine  
N/Alternate names: aspartyl (asparaginyl) beta-hydroxylase  
C/Species: Bos primigenius taurus (cattle)  
C/Date: 31-Dec-1993 #sequence\_revision 10-Feb-1995 #text\_change 11-Jun-1999  
C/Accession: A42969; A39470; B39470; C39470; S27948  
R/Jia, S.; Vandusen, W.J.; Diehl, R.E.; Kohl, N.E.; Dixon, R.A.; Elliston, K.O.; Ster  
J. Biol. Chem. 267, 14322-14327, 1992  
A/Title: cDNA cloning and expression of bovine aspartyl (asparaginyl) beta-hydroxylas  
A/Reference number: A42969; MUID:92332546; PMID:1378441  
A/Accession: A42969  
A/Molecule type: mRNA  
A/Residues: 1-754 <JIA>  
A/Cross-references: EMBL:M91213; NID:g162693; PIDN:AAA03563.1; PID:g162694  
A/Experimental source: brain  
A/Note: Sequence extracted from NCBI backbone (NCBI:P108534)  
R/Wang, O.; Vandusen, W.J.; Petroski, C.J.; Garsky, V.M.; Stern, A.M.; Friedman, P.A.  
J. Biol. Chem. 266, 14004-14010, 1991  
A/Title: Bovine liver aspartyl beta-hydroxylase. Purification and characterization.  
A/Reference number: A39470; MUID:91310689; PMID:1856229  
A/Accession: A39470  
A/Molecule type: protein  
A/Residues: 289-328 <WAN>  
A/Accession: B39470  
A/Molecule type: protein  
A/Residues: 615, 'X', 617-630, 'XX', 633-634, 'X', 636, 'XX', 639-641 <WA2>  
A/Accession: C39470  
A/Molecule type: protein  
A/Residues: 311-347, 'X', 349, 'X', 351-373, 'X', 375-379, 'X', 381-382 <WA3>  
C/Comment: This enzyme uses ferrous iron as a cofactor, and while beta-hydroxylating  
C/Keywords: aspartic acid and asparagine residues in the EGF homology domain of certai  
C/Superfamily: peptidase-aspartate beta-dioxygenase; tetrairicopeptide repeat homology  
F:2-56/Domain: intracellular #status predicted <IRM>  
F:57-78/Domain: transmembrane #status predicted <TRM>  
F:289-754/Product: peptide-aspartate beta-dioxygenase, 56K form #status predicted <S6  
F:311-754/Product: peptide-aspartate beta-dioxygenase, 52K form #status predicted <S2  
F:337-370/Domain: tetrairicopeptide repeat homology <TT1>

F:371-404/Domain: tetratricopeptide repeat homology <TRP>  
E:13,96,466,702/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 43.1%; Score 43.5; DB 1; Length 754;  
Best Local Similarity 55.6%; Pred. No. 93;

Matches 10; Conservative 3; Mismatches 4; Indels 1; Gaps 1;

OY 1 DLEETIANSKDFP-NMSE 17  
DB 82 DYEETLAKAKDFRYNLSE 99

## RESULT 14

S65571

Pattern formation protein GNOM - Arabidopsis thaliana

N:Alternate names: EMB30 protein

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 07-May-1999

C:Accession: S65571; S65572

C:Accession: S65571; S65572

A:Reference number: S65571; MUID:96204508; PMID:8628228

A:Reference number: S65571; MUID:96204508; PMID:8628228

A:Accession: S65571

A:Molecule type: DNA

A:Residues: 1-1451 <BUS>

A:Cross-References: EMBL:U36433; NID:g1209632; PID:g1209633

A:Accession: S65572

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-110, 'T', 112-866, 'G', 868-1451 <BMW>

A:Cross-References: EMBL:U36432; NID:g1209630; PID:g1209631

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1995

C:Genetics:

A:Gene: GNOM: EMB30

A:introns: 246/3

Query Match 43.1%; Score 43.5; DB 2; Length 1451;  
Best Local Similarity 63.2%; Pred. No. 2e+02;

Matches 12; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

OY 1 DLEETIANS-KDFPMSET 18  
DB 1302 DLEETIAGSQKDYRNMEGT 1320

## RESULT 15

T29144

Partial CDS - Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 31-Jan-2000

C:Accession: T29144

R:Pauley, A.; Galtung, S.

A:Description: The sequence of C. elegans cosmid K11C4.

A:Reference number: Z20577

A:Accession: T29144

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-5107 <PAN>

A:Cross-References: EMBL:U64854; PIDN:AA18318.1; GSPDB:GN00023; CESP:unc-68

A:Experimental source: strain Bristol N2; clone K11C4

C:Genetics:

A:Gene: CESP:unc-68

A:Map position: 5

A:introns: 27/1; 64/3; 92/3; 127/1; 158/2; 1222/2; 1300/2; 1347/2; 1391/1; 1419/3; 1517/

C:Superfamily: Ryanodine receptor; transcription initiation factor sigma region 1 homolog

Query Match 43.1%; Score 43.5; DB 2; Length 5107;  
Best Local Similarity 43.5%; Pred. No. 8.4e+02;  
Matches 10; Conservative 4; Mismatches 6; Indels 3; Gaps 1;

OY 1 DLEETIANS---KDFPMSETDL 20  
DB 3377 DLEETIANNNTMTSDVPNVYDVL 3399

Search completed: October 1, 2003, 10:04:43  
Job time: 15.8235 secs

**THIS PAGE BLANK (USPTO)**



FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.  
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.  
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.  
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.  
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80B9BE75 CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2226;  
Best Local Similarity 100.0%; Pred. No. 1,1e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETANSKDFPNMSETDL 20  
DB 931 DLEETANSKDFPNMSETDL 950

## RESULT 2

POLG\_HPAV8 STANDARD: PRT: 2226 AA.

AC P26582; 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DE 15-SEP-2003 (Rel. 42, Last annotation update)

OS Hepatitis A virus (strain 18f).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OX NCBI\_TaxID=12096;

RN [1]

RP MEDLINE-91162758; PubMed-1705995;

RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,

RT "Antigenic and genetic variation in cytopathic hepatitis A virus

RL J. Virol. 65:2056-2065(1991).

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +

CC (RNA)(N).

CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,

CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,

CC VP3, AND VP4.

CC -1- FUNCTION: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: M59808; AAA45467.1; -

CC PDB: 1OAT; 15-MAY-00.

CC INTERPRO: IPR000605; RNA\_helicase.

CC INTERPRO: IPR007095; RNA\_pol\_DS\_PS.

CC INTERPRO: IPR001205; RNA\_pol\_P3D.

CC INTERPRO: IPR007094; RNA\_pol\_PSVIT.

CC Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.

CC Pfam: PF00910; RNA\_helicase; 1.

CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease; 3D-structure.

FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.  
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.  
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.  
SQ SEQUENCE 2226 AA; 251292 MW; 24964A63396C8D6B CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2226;  
Best Local Similarity 100.0%; Pred. No. 1,1e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETANSKDFPNMSETDL 20  
DB 931 DLEETANSKDFPNMSETDL 950

## RESULT 3

POLG\_HPAV8 STANDARD: PRT: 2227 AA.

AC P06617; P06443; 081082;

DT 01-AUG-1988 (Rel. 08, Created)

DE 15-SEP-2003 (Rel. 42, Last sequence update)

OS Hepatitis A virus (strain HM-175).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OX NCBI\_TaxID=12098;

RN [1]

RP MEDLINE-87061253; PubMed-3023706;

RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,

RT "Complete nucleotide sequence of wild-type hepatitis A virus:

RL J. Virol. 61:50-59(1987).

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +

CC (RNA)(N).

CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,

CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,

CC VP3, AND VP4.

CC -1- FUNCTION: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT

CC SHOWN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: M59808; AAA45467.1; -

CC PDB: 1OAT; 15-MAY-00.

CC INTERPRO: IPR000605; RNA\_helicase.

CC INTERPRO: IPR007095; RNA\_pol\_DS\_PS.

CC INTERPRO: IPR001205; RNA\_pol\_P3D.

CC INTERPRO: IPR007094; RNA\_pol\_PSVIT.

CC Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.

CC Pfam: PF00910; RNA\_helicase; 1.

CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease; 3D-structure.

or send an email to [license@lsb.slb.ch](mailto:license@lsb.slb.ch)).

CC EMBL; M14114; AAA45475.1; -  
 DR EMBL; M14707; AAA45465.1; -  
 DR EMBL; M14707; AAA45466.1; ALT\_INIT.  
 DR EMBL; M16632; AAA45471.1; -  
 DR PIR; A25981; GNNYWK.  
 DR PIR; A94149; GNNYWK.  
 DR PDB; 1HAV; 23-DEC-96.  
 DR MEROPS; C03.005; -  
 DR InterPro; IPR004004; Calic1\_pol\_hel.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_PSVir.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol.1.  
 DR Pfam; PF00910; RNA\_helicase.1.  
 DR PRINTS; PR00918; CALICVIRUSNS.  
 KM Polyprotein; Coat protein; Transferase;  
 RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.  
 CHAIN 1 23  
 FT CHAIN 24 245 COAT PROTEIN VP4 (P1A).  
 FT CHAIN 246 491 COAT PROTEIN VP2 (P1B).  
 FT CHAIN 482 836 COAT PROTEIN VP3 (P1C).  
 FT CHAIN 837 980 COAT PROTEIN VP1 (P1D).  
 FT CHAIN 981 1087 CORE PROTEIN P2A.  
 FT CHAIN 1088 1422 CORE PROTEIN P2B.  
 FT CHAIN 1423 1496 CORE PROTEIN P2C.  
 FT CHAIN 1497 1519 PROBABLE PROTEIN P3A.  
 FT CHAIN 1520 1738 PROBABLE PROTEIN P3B.  
 FT CHAIN 1739 2227 PROBABLE PROTEIN P3C.  
 FT CHAIN 2227 2227 RNA-DIRECTED POLYMERASE 3D.  
 FT CHAIN 77 77 K -> R (IN ATTENUATED STRAIN).  
 FT CHAIN 764 764 E -> V (IN ATTENUATED STRAIN).  
 FT CHAIN 821 821 N -> S (IN ATTENUATED STRAIN).  
 FT CHAIN 1052 1052 A -> V (IN ATTENUATED STRAIN).  
 FT CHAIN 1062 1062 G -> A (IN ATTENUATED STRAIN).  
 FT CHAIN 1118 1118 K -> M (IN ATTENUATED STRAIN).  
 FT CHAIN 1151 1151 E -> K (IN ATTENUATED STRAIN).  
 FT CHAIN 1163 1163 F -> S (IN ATTENUATED STRAIN).  
 FT CHAIN 1277 1277 V -> I (IN ATTENUATED STRAIN).  
 FT CHAIN 1500 1500 H -> Y (IN ATTENUATED STRAIN).  
 FT CHAIN 1805 1805 D -> N (IN ATTENUATED STRAIN).  
 FT CHAIN 1930 1930 S -> T (IN ATTENUATED STRAIN).  
 FT CHAIN 2227 2227 S -> T (IN ATTENUATED STRAIN).  
 SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AE8740A6 CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEETIANSKDFPNMSETDL 20  
 ||||||||||||||||  
 Db 931 DLEETIANSKDFPNMSETDL 950

RESULT 4  
 POLG\_HPAVL STANDARD; PRT; 2227 AA.  
 ID POLG\_HPAVL  
 AC P06441;  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
 DE Hepatitis A virus (strain LA).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Hepatovirus.  
 OC Hepatovirus.  
 OC NCBI\_TaxID=12099;  
 OX NCBI\_TaxID=12099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85190549; PubMed=2986127;  
 RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A., Merryweather J., Van Nest G., Dina D.;

\*Primary structure and gene organization of human hepatitis A virus.\*;  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +  
 (RNA)(N).  
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
 EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
 VP3, AND VP4.  
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.lsb.slb.ch/announce/>  
 or send an email to [license@lsb.slb.ch](mailto:license@lsb.slb.ch)).

CC EMBL; K02990; AAA45472.1; -  
 DR PIR; A03903; GNNYWK.  
 DR MEROPS; C03.005; -  
 DR InterPro; IPR004004; Calic1\_pol\_hel.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_PSVir.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol.1.  
 DR Pfam; PF00910; RNA\_helicase.1.  
 DR PRINTS; PR00918; CALICVIRUSNS.  
 KM Polyprotein; Coat protein; Transferase;  
 RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
 CHAIN 1 23  
 FT CHAIN 24 245 COAT PROTEIN VP4 (P1A).  
 FT CHAIN 246 491 COAT PROTEIN VP2 (P1B).  
 FT CHAIN 492 836 COAT PROTEIN VP3 (P1C).  
 FT CHAIN 837 980 COAT PROTEIN VP1 (P1D).  
 FT CHAIN 981 1076 CORE PROTEIN P2A.  
 FT CHAIN 1077 1422 CORE PROTEIN P2B.  
 FT CHAIN 1423 1484 CORE PROTEIN P2C.  
 FT CHAIN 1485 1507 PROBABLE PROTEIN P3A.  
 FT CHAIN 1508 1678 PROBABLE PROTEIN P3B.  
 FT CHAIN 1679 2227 PROBABLE PROTEIN P3C.  
 FT CHAIN 2227 2227 RNA-DIRECTED POLYMERASE P3D.  
 SQ SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 101; DB 1; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-07;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEETIANSKDFPNMSETDL 20  
 ||||||||||||||||  
 Db 931 DLEETIANSKDFPNMSETDL 950

RESULT 5  
 POLG\_HPAVL STANDARD; PRT; 2227 AA.  
 ID POLG\_HPAVL  
 AC P13901; Q81083; Q81084; Q81085; Q81086; Q81087; Q81088; Q81089;  
 AC Q81090; Q81091; Q81092; Q81093;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
 DE Hepatitis A virus (strain HBB).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Hepatovirus.  
 OC Hepatovirus.  
 OC NCBI\_TaxID=12100;  
 OX NCBI\_TaxID=12100;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88045071; PubMed=2823500;  
 RA Paul A.V., Tada H., der Helm K., Wessel T., Klehn R., Wimmer E., Deinhardt F.;

```

RT      "The entire nucleotide sequence of the genome of human hepatitis A
RT      virus (isolate MBF).";
RL      Virus Res. 8:153-171(1987).
CC      -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC      (RNA)(N).
CC      -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC      EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC      VP3, AND VP4.
CC      -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC      -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT
CC      WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.
CC      -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M20273; AAA5474.1; -.
DR      MEROPS; C03.005; -.
DR      InterPro; IPR000605; RNA_helicase.
DR      InterPro; IPR007095; RNA_pol_DS_PS.
DR      InterPro; IPR001205; RNA_pol_P3D.
DR      InterPro; IPR007094; RNA_pol_PSVtr.
DR      Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR      Pfam; PF00910; RNA_helicase; 1.
KM      Polypeptide: Coat protein: Core protein: Transferase:
KM      RNA-directed RNA polymerase: Hydroxylase; Thiol protease.
FT      CHAIN
FT      1
FT      CHAIN
FT      24
FT      CHAIN
FT      245
FT      CHAIN
FT      466
FT      CHAIN
FT      491
FT      CHAIN
FT      492
FT      CHAIN
FT      837
FT      CHAIN
FT      837
FT      CHAIN
FT      980
FT      CHAIN
FT      1087
FT      CHAIN
FT      1088
FT      CHAIN
FT      1422
FT      CHAIN
FT      1423
FT      CHAIN
FT      1496
FT      CHAIN
FT      1497
FT      CHAIN
FT      1519
FT      CHAIN
FT      1738
FT      CHAIN
FT      1738
FT      CHAIN
FT      1739
FT      CHAIN
FT      2227
SQ      SEQUENCE
SQ      2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;

Query Match      100.0%; Score 101; DB 1; Length 2227;
Best Local Similarity 100.0%; Pident. No. 1; E-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEETIANSKDFPNMSETDL 20
|||||
931 DLEETIANSKDFPNMSETDL 950

RESULT 6
POLG_HPAVS
ID POLG_HPAVS STANDARD; PRT; 2230 AA.
AC P14553;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Simian hepatitis A virus (strain AGM-27).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC P14553;
OX NCBI_TaxID=12102;
[1]
RX SEQUENCE FROM N.A.
RA MEDLINE=913111420; PubMed=1649901;
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ricehurst J.R.,
RA Purcell R.H.;
RA "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome

```

```

RT structure and growth in cell culture with other HAV strains."
RL J. Gen. Virol. 72:1677-1683(1991).
RN [2]
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE=89232168; PubMed=2541023;
RA Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
RA Sverdlov E.D., Chizhikov V.E., Bilnov V.M., Vasilenko S.K.;
RT "Variations in genome fragments coding for RNA polymerase in human
RL and simian hepatitis A viruses."
CC FEBS Lett. 247:425-428(1989).
CC -1 CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA)n.
CC
CC -1 SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1 PM1: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-stb.ch/announce/
CC or send an email to licenses@isb-stb.ch).
CC -----
DR EMBL; D00924; BA00766.1; -;
DR EMBL; X15461; CA343490.1; -;
DR PIR; A30470; GNNYSA.
DR MEROPS; C03.005; -;
DR InterPro: IPR004004; Calic1.pol.hel.
DR InterPro: IPR0000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_psv1t.
DR Pfam; PF00680; RNA_dep.RNA.pol.1.
DR Pfam; PF00910; RNA_helicase.1.
DR PRINTS; PR00918; CALICYRUSNS.
DR PolyProtein; Coat protein; Core protein; Transferase;
RW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 27
FT CHAIN 28 249
FT CHAIN 250 495
FT CHAIN 496 795
FT CHAIN 796 984
FT CHAIN 985 1091
FT CHAIN 1092 1426
FT CHAIN 1427 1498
FT CHAIN 1499 1521
FT CHAIN 1522 1741
FT CHAIN 1742 2230
FT CHAIN 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;
SQ SEQUENCE
QY Query Match 97.0%; Score 98; DB 1; Length 2230;
Db Best Local Similarity 95.0%; Pred. No. 3.3e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLEETIAANSKDFPNMSSETDL 20
QY |||||:|||||
Db 935 DLEETIAANSKDFPNMSSETDL 954
RESULT 7
POLG_HPAV2 STANDARD; PRT; 2226 AA.
AC P26580:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
DE P3D (EC 2.7.7.48)].
OS Hepatitis A virus (strain 24a).

```



```

CC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
RA NCBI_TaxID=12094.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91162758; PubMed=1705995;
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,
RA Croomeans T., Jansen R.W.;
RT "Antigenic and genetic variation in cytopathic hepatitis A virus
RT variants arising during persistent infection: evidence for genetic
RT recombination."
RL J. Virol. 65:2056-2065(1991).
CC -1 CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC {N}(N)
CC -1 SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1 PWM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1 SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isd-sdb.ch/announce/
CC or send an email to license@sdb-sdb.ch).
CC -----
DR EMBL; M59810; AAA45468.1; -.
DR MEROPS; C03.005; -.
DR InterPro; IPR004A004; Calicl_pol_hel.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_P5.
DR InterPro; IPR001205; RNA_pol_DS_P3.
DR InterPro; IPR007094; RNA_pol_PSVlr.
DR Pfam; PF00680; RNA_dep_RNA_pol_1.
DR Pfam; PF00910; RNA_helicase_1.
DR PRINTS; PR00918; CALICVIRUSNS.
KM Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.
FT CHAIN 1 23
FT CHAIN 24 245
FT CHAIN 246 491
FT CHAIN 492 794
FT CHAIN 795 900
FT CHAIN 901 1087
FT CHAIN 1088 1422
FT CHAIN 1423 1495
FT CHAIN 1496 1518
FT CHAIN 1519 1737
FT CHAIN 1738 2226
SO SEQUENCE 2226 AA; 251152 MW; 6CD85A91DB5B4E2BF CRC64;

Query Match 96.0%; Score 97; DB 1; Length 2226;
Best Local Similarity 95.0%; Pred. No. 4.8e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 DLEEIAANSKDFPNMSETDL 20
| | | | | | | | | | | | | | | | | | | | | |
Db 931 DLEEIAANSKEFPNMSETDL 950

RESULT 8
Y041_THEAC STANDARD; PRT: 194 AA.
AC P57674;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
GN Hypothetical protein Ts0041.
OS Thermoplasma acidophilum.
OC Archaeae; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
```

CC	Thermoplasmataceae; Thermoplasma.
OX	MGBL_TaxID=2303;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-DSM 1728;
RX	MEDLINE=20479972; PubMed=11029001;
RA	Ruepp A., Gräml W., Santos-Martinez M.-L., Koretke K.R., Volker C.,
RT	Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumelster W.;
RL	"The genome sequence of the thermoacidophilic scavenger Thermoplasma
CC	acidophilum".
CC	Nature 407:508-513 (2000).
CC	-1- SIMILARITY: BELONGS TO THE DP60129 FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; AL445063; CAC11190.1; -
DR	HAMAP; MF_00265; -; 1.
DR	IoterpPro; IPR0002851; DUF133.
DR	IoterpPro; IPR002716; PIN.
DR	IoterpPro; IPR006596; PING.
DR	Pfam; PF01850; PIN; 1.
DR	ProDom; PD013236; DUF133; 1.
DR	SMART; SM00670; PING; 1.
KW	Hypothetical protein; Complete proteome.
SQ	SEQUENCE 194 AA; 21854 MW; 6E7CCDD844F9FFAF CRC64;
OY	Query Match 45.5%; Score 46; DB 1; Length 194;
Db	Best Local Similarity 50.0%; Pred. No. 4.1;
	Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
OY	2 LEETANSKDEPPNMSETD 19
	::::: ::: -:-
Db	92 VETPAKTGDLMLNSQTD 109
RESULT 9	
GAG-SMRVH	STANDARD; PRT; 740 AA.
ID	GAG-SMRVH
AC	P21411;
DT	01-MAY-1991 (Rel. 18, Created)
DT	01-MAY-1991 (Rel. 18, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	GAG polypeptide [Contains: Core protein P19; Core protein P16;
DE	Probable core protein P35; Probable core protein P10].
GN	GAG.
OS	Squirrel monkey retrovirus (SMRV-H) (SMRV-HLB)
OC	Viruses; Retroviral viruses; Retroviridae; Betaretrovirus.
OX	MGBL_TaxID=11836;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=89073750; PubMed=3201749;
RA	Oda T., Ikeda S., Watanabe S., Hatsushika M., Akiyama K.,
RT	Mitsunobu F.;
RL	"Molecular cloning, complete nucleotide sequence, and gene structure
CC	of the provirus genome of a retrovirus produced in a human
CC	lymphoblastoid cell line.";
CC	Virology 167:468-476 (1988).
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
EMBL; M23385; AAA66451.1; -	



DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Aspartyl-asparaginyl beta-hydroxylase (EC 1.14.11.16) (Aspartate beta-  
 DE hydroxylase) (ASP beta-hydroxylase) (Peptide-aspartate beta-  
 DE dihydroxylase).  
 GN ASPH.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 CX NCBI\_TaxID:9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain, and Liver;  
 RX MEDLINE=92332546; PubMed=1378641;  
 RA Jia S., Vandenusen W.J., Dieml R.E., Kohl N.E., Dixon R.A.F.,  
 RA Elliston K.O., Stern A.M., Friedman P.A.;  
 RT "cDNA Cloning and expression of bovine aspartyl (asparaginyl) beta-  
 RT hydroxylase.";  
 RL J. Biol. Chem. 267:14322-14327(1992).  
 [2]  
 SEQUENCE OF 289-385 AND 615-641.  
 TISSUE=Liver;  
 RX MEDLINE=91310689; PubMed=1856229;  
 RA Wang Q., Vandenusen W.J., Petroski C.J., Garaky V.M., Stern A.M.,  
 RA Friedman P.A.;  
 RT "Bovine liver aspartyl beta-hydroxylase. Purification and  
 RT characterization.";  
 RL J. Biol. Chem. 266:14004-14010(1991).  
 CC -1- FUNCTION: SPECIFICALLY HYDROXYLATES AN ASP OR ASN RESIDUE IN  
 CC CERTAIN EPIDERMAL GROWTH FACTOR-LIKE (EGF) DOMAINS OF A NUMBER OF  
 CC PROTEINS.  
 CC -1- CATALYTIC ACTIVITY: Peptide L-aspartate + 2-oxoglutarate + O(2) =  
 CC peptide 3-hydroxy-L-aspartate + succinate + CO(2).  
 CC -1- COFACTOR: IRON.  
 CC -1- SUBUNIT: Monomer.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic  
 CC reticulum.  
 CC -1- PMW: MIGHT BE PROCESSED TO THE 56 kDa (AA 289-754) OR 52 kDa (AA  
 CC 311-754) FORMS IN THE LUMEN OF THE ENDOPLASMIC RETICULUM.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL, M91213; AAA03563.1; -.  
 CC PIR: A42969; BABOH.  
 CC InterPro: IPR001440; TPR.  
 DR Pfam: PF05279; Asp-B-Hydro\_N; 1.  
 DR Pfam: PF05118; Asp.Arg\_Hydrox; 1.  
 KW Oxidoreductase; Dioxigenase; Iron; Transmembrane; Signal-anchor;  
 KW Endoplasmic reticulum.  
 FT DOMAIN 1 57 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 58 78 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT (POTENTIAL).  
 FT LDMENAL (POTENTIAL).  
 FT DOMAIN 79 754 POLY-GLY.  
 FT DOMAIN 14 21 POLY-SER.  
 FT 318 328 POLY-LYS.  
 FT CARBOHYD 96 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 702 702 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 754 AA: 84998 MW: 369592Alr0B558C6 CRC64;  
 Query Match 43.1%; Score 43.5; DB 1; Length 754;  
 Best Local Similarity 55.6%; Pred. No. 42;  
 Matches 10; Conservative 3; Mismatches 4; Indels 1; Gaps 1

Db 82 DYEVLAKANDERYNLSE 99

```
RESULT 12
ID EM30_ARATH STANDARD: PRT: 1451 AA.
AC 042510: 038883:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Pattern formation protein EM30.
GN EM30 OR GNOM OR ATG13980 OR F7A19.7 OR F16A14.20.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustoids II; Brassicales; Brassicaceae; Arabidopsis.
OC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Wassilewskija, and cv. Columbia;
RX MEDLINE=944291195; PubMed=8020095;
RA Shevell D.E., Leu W.-M., Gillmor C.S., Xia G., Feldmann K.A.,
RA Chua N.-H.;
RT "EM30 is essential for normal cell division, cell expansion, and
RT cell adhesion in Arabidopsis and encodes a protein that has
RT similarity to Sec7."
RL Cell 77:1051-1062(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia, and cv. CSR-1;
RX MEDLINE=96204508; PubMed=8628228;
RA Busch M., Mayer U., Uebersch G.;
RT "Molecular analysis of the Arabidopsis pattern formation of gene
RT GNOM: gene structure and intergenic complementation."
RL Mol. Gen. Genet. 250:681-691(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altar J.H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewart K.,
RA Dunn P., Elgu J., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Guller J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetskaia I., Kurtz D.B., Kvan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Malti R., Marshall A.,
RA Maltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA San H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana."
RL Nature 408:816-820(2000).
CC -I- FUNCTION: MAY PERFORM A FUNCTION THAT AFFECTS CELL EXPANSION, THE
CC ORIENTATION OF THE PLANE OF CELL DIVISION, THE NUMBER OF CELL
CC DIVISIONS, AND CELL ADHESION THROUGHOUT PLANT DEVELOPMENT.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -I- TISSUE SPECIFICITY: STEMS, LEAVES, FLOWERS, SILIQUES, FLORAL
CC INFLORESCENCE AND ROOTS.
CC -I- SIMILARITY: Contains 1 Sec7 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

CC EMBL: U56140; AAB01205.1; -  
DR EMBL: U56141; AAB01206.1; -  
DR EMBL: U36432; AAA91150.1; -  
DR EMBL: U36433; AAA91151.1; -  
DR EMBL: AC007576; AAD39284.1; -  
DR EMBL: AC068197; AAF9403.1; -  
DR PIR: S65571; S65571.  
DR HSP: Q99418; IPBV.  
DR InterPro: IPR000904; Sec7.  
DR Pfam: PF01369; Sec7; 1.  
DR SMART: SM00222; Sec7; 1.  
DR PROSITE: PS50190; SEC7; 1.  
KW Cell adhesion.  
FT DOMAIN 357 752 SEC7.  
FT MOTIF 658 658 E->K: EMB30-1.  
FT CONFLICT 111 111 T->I (IN REF. 2; AAA91150).  
FT CONFLICT 867 867 A->G (IN REF. 2; AAA91150).  
SQ SEQUENCE 1451 AA; 162618 MW; 666E21C74E426996 CRC64;  
Query Match 43.18; Score 43.5; DB 1; Length 1451;  
Best Local Similarity 63.28; Pred. No. 83;  
Matches 12; Conservative 1; Mismatches 5; Indels 1; Gaps 1;  
QY 1 DLEBIANS-KDFPNSMSET 18  
DB 1302 DLEIAGSQRKDYRMEGT 1320  
ID UPPE\_AQUAE STANDARD; PRT; 231 AA.  
AC 067291;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Undecaprenyl pyrophosphate synthetase (EC 2.5.1.31) (UPP synthetase)  
DE (di-trans-poly-cis-decaprenylcistransfetae) (Undecaprenyl diphosphate  
DE synthase) (UDS).  
GN UPPE OR AQ\_1248.  
OS Aquifex aeolicus.  
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.  
OX NCBI\_TaxID=63363;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VFS;  
RX MEDLINE=9819666; PubMed=9537320;  
RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
Graham D.E., Overbeek R., Snead M.A., Keller M., Auslay M., Huber R.,  
Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
"The complete genome of the hyperthermophilic bacterium Aquifex  
aeolicus.";  
RT Nature 392:353-358(1998).  
CC -1- FUNCTION: GENERATES UNDECAPRENYL PYROPHOSPHATE (UPP) FROM  
ISOPENTENYL PYROPHOSPHATE (IPP). UPP IS THE PRECURSOR OF THE  
CARRIER LIPID FOR PEPTIDOGLYCAN SYNTHESIS (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Di-trans-poly-cis-decaprenyl diphosphate +  
isopentenyl diphosphate -> diphosphate + di-trans-poly-cis-  
undecaprenyl diphosphate.  
CC -1- SIMILARITY: BELONGS TO THE UPP SYNTHETASE FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AE000730; AAC07254.1; -  
DR PIR: H70407; H70407.  
DR InterPro: IPR001441; UPP\_synth.  
DR Pfam: PF01255; UPP\_synthetase; 1.

DR PRODOM: PD003461; UPP\_synth; 1.  
DR TIGRFAMs: TIGR00055; upps; 1.  
DR PROSITE: PS01066; UPP\_SYNTHETASE; 1.  
KW Transferase; Cell division; Cell wall; Peptidoglycan synthesis;  
KW Complete proteome.  
SQ SEQUENCE 231 AA; 27148 MW; 52B9DE0A44280AE CRC64;  
Query Match 42.68; Score 43; DB 1; Length 231;  
Best Local Similarity 50.08; Pred. No. 15;  
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;  
QY 2 LEEIANSKDFPNN 15  
DB 112 MEELESDSKDFKNL 125  
ID 40MT\_COPUA STANDARD; PRT; 350 AA.  
AC 09LELS;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE 3'-hydroxy-N-methyl-(S)-coclaurine 4'-O-methyltransferase  
DE (EC 2.1.1.116) (S-adenosyl-L-methionine:3'-hydroxy-N-methylcoclaurine  
DE 4'-O-methyltransferase) (4'-OMT).  
OS Coplis japonica (Japanese goldthread).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;  
OC Ranunculaceae; Coplis.  
OX NCBI\_TaxID=3442;  
RN [1]  
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RX MEDLINE=20390108; PubMed=10811648;  
RA Moshige T., Tsujita T., Yamada Y., Sato F.;  
"Molecular characterization of the S-adenosyl-L-methionine:  
3'-hydroxy-N-methylcoclaurine 4'-O-methyltransferase involved in  
RT isequinolone alkaloid biosynthesis in Coplis japonica.";  
RL J. Biol. Chem. 275:23398-23405(2000).  
CC -1- FUNCTION: CATALYZES THE TRANSFER OF THE METHYL GROUP TO THE 4'-  
HYDROXYL GROUP OF 3'-HYDROXY-N-METHYLCOCCLAURINE TO FORM  
CC RETICULINE.  
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 3'-hydroxy-N-methyl-  
CC (S)-coclaurine -> S-adenosyl-L-homocysteine + (S)-reticuline.  
CC -1- PATHWAY: CARRIES OUT A STEP IN THE CONVERSION OF 3'-HYDROXY-N-  
CC METHYLCOCCLAURINE TO RETICULINE, AN IMPORTANT INTERMEDIATE IN  
CC SYNTHESIZING ISOQUINOLINE ALKALOIDS.  
CC -1- SUBUNIT: Homodimer.  
CC -1- SIMILARITY: BELONGS TO THE METHYLTRANSFERASE SUPERFAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D29812; BAB08005.1; -  
DR InterPro: IPR001601; Methyltransf.  
DR InterPro: IPR001077; O-Methyltransf.  
DR Pfam: PF00891; Methyltransf\_2; 1.  
KW Transferase; Methyltransferase.  
SQ SEQUENCE 350 AA; 38775 MW; 547835EBCDEF9182 CRC64;  
Query Match 42.68; Score 43; DB 1; Length 350;  
Best Local Similarity 53.38; Pred. No. 22;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
QY 1 DLEBIANSKDFPNN 15  
DB 219 DLEPHVANSYDLRNI 233

RESULT 15  
 RPOB\_CAMJE STANDARD; PRT; 1378 AA.  
 ID RPOB\_CAMJE 046124; 09P131;  
 AC 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase  
 beta chain) (RNA polymerase beta subunit).  
 GN RPOB OR CJO478.  
 OS Campylobacter jejuni.  
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 OC Campylobacteraceae; Campylobacter.  
 OX NCBI\_TaxID=197;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NOTC 11168;  
 RX MEDLINE=20150912; PubMed=10688204;  
 RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,  
 Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,  
 Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
 Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
 Whitehead S., Barrett B.G.,  
 RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
 reveals hypervariable sequences.";  
 RL Nature 403:665-668(2000).  
 RN [2]  
 RP SEQUENCE OF 338-1031 FROM N.A.  
 RX MEDLINE=96084944; PubMed=7489896;  
 RA Bustamante V.H., Puente J.L., Sanchez-Lopez F., Bobadilla M.,  
 Calva E.;  
 RT "Identification of Campylobacter jejuni and C.coli using the rpoB  
 gene and a cryptic DNA fragment from C.jejuni.";  
 RL Gene 165:1-8(1995).  
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
 SUBSTRATES.  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 {RNA}(N).  
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE  
 ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1  
 BETA' CHAIN.  
 CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AL139075; CAB75116.1; -  
 DR EMBL: X77304; CA54509.1; -  
 DR PIR: A81393; A81393.  
 DR PIR: S41868; S41868.  
 DR HSSP: Q9KMU7; 1HOM.  
 DR InterPro: IPR001572; RNA\_pol.B.  
 DR Pfam: PF04563; RNA\_pol\_Rpb2\_1; 1.  
 DR Pfam: PF04561; RNA\_pol\_Rpb2\_2; 1.  
 DR Pfam: PF04565; RNA\_pol\_Rpb2\_3; 1.  
 DR Pfam: PF00562; RNA\_pol\_Rpb2\_6; 1.  
 DR Pfam: PF04560; RNA\_pol\_Rpb2\_7; 1.  
 DR PROSITE: PS01166; RNA\_POL\_BETA; 1.  
 DR Transfaser: Transcription; DNA-directed RNA polymerase;  
 KW Complete proteome.  
 KM CONFLICT 338 347 NDIANGVDA -> MTWLMALMP (IN REF. 2).  
 FT CONFLICT 558 558 A -> R (IN REF. 2).  
 FT CONFLICT 671 671 C -> S (IN REF. 2).  
 FT CONFLICT 691 691 A -> R (IN REF. 2).  
 SO SEQUENCE 1378 AA; 155915 MW; AB7467C305028EB5 CRC64;

Query Match 41.18; Score 41.5; DB 1; Length 1378;  
 Best Local Similarity 47.48; Pred. No. 1.6e+02;  
 Matches 9; Conservative 4; Mismatches 3; Indels 3; Gaps 1;  
 QY 2 LEEIANSKDFPMSETDL 20  
 Db 858 IEEI---TKDIPNVKEDV 873

Search completed: October 1, 2003, 09:57:52  
 Job time : 8.58824 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 33.1765 seconds  
(without alignments)  
155.564 Million cell updates/sec

Title: US-09-171-432A-45

Perfect score: 101  
Sequence: 1 DLEETANSKDFPNMSETDL 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Minimum number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: SP\_archaea:\*
- 2: SP\_bacteria:\*
- 3: SP\_fungi:\*
- 4: SP\_human:\*
- 5: SP\_invertebrate:\*
- 6: SP\_mammal:\*
- 7: SP\_mmc:\*
- 8: SP\_organelle:\*
- 9: SP\_phage:\*
- 10: SP\_plant:\*
- 11: SP\_rodent:\*
- 12: SP\_virus:\*
- 13: SP\_vertebrate:\*
- 14: SP\_unclassified:\*
- 15: SP\_virus:\*
- 16: SP\_bacteriaph:\*
- 17: SP\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	251	12	Q9ENQ4
2	101	100.0	251	12	Q9ENN2
3	101	100.0	251	12	Q9ENP2
4	101	100.0	251	12	Q9ENQ6
5	101	100.0	251	12	Q9ENR1
6	101	100.0	251	12	Q9ENP1
7	101	100.0	251	12	Q9ENQ1
8	101	100.0	251	12	Q9ENQ9
9	101	100.0	251	12	Q9ENR4
10	101	100.0	251	12	Q9ENP5
11	101	100.0	251	12	Q9ENP7
12	101	100.0	251	12	Q9ENQ5
13	101	100.0	251	12	Q9ENP9
14	101	100.0	251	12	Q9ENN6
15	101	100.0	251	12	Q9ENN5
16	101	100.0	251	12	Q9ENN7

17	101	100.0	1124	12	Q84780	Q84780 hepatitis a
18	101	100.0	1161	12	Q05794	Q05794 hepatitis a
19	101	100.0	2216	12	Q9WMA2	Q9WMA2 hepatitis a
20	101	100.0	2218	12	Q67824	Q67824 hepatitis a
21	101	100.0	2225	12	Q9DLJ2	Q9DLJ2 hepatitis a
22	101	100.0	2227	12	Q9WMA0	Q9WMA0 hepatitis a
23	101	100.0	2227	12	Q9WMA3	Q9WMA3 hepatitis a
24	101	100.0	2227	12	Q67825	Q67825 hepatitis a
25	101	100.0	2227	12	Q9WMA1	Q9WMA1 hepatitis a
26	101	100.0	2227	12	Q67826	Q67826 hepatitis a
27	101	100.0	2227	12	Q8VON6	Q8VON6 hepatitis a
28	101	100.0	2227	12	Q9IFH5	Q9IFH5 hepatitis a
29	101	100.0	2227	12	Q9WMA4	Q9WMA4 hepatitis a
30	95	94.1	251	12	Q9ENN9	Q9ENN9 hepatitis a
31	95	94.1	251	12	Q9ENN8	Q9ENN8 hepatitis a
32	95	94.1	251	12	Q9ENN3	Q9ENN3 hepatitis a
33	95	94.1	251	12	Q9ENN7	Q9ENN7 hepatitis a
34	95	94.1	251	12	Q9ENP0	Q9ENP0 hepatitis a
35	95	94.1	251	12	Q9ENP4	Q9ENP4 hepatitis a
36	95	94.1	251	12	Q9ENQ3	Q9ENQ3 hepatitis a
37	95	94.1	251	12	Q9ENP6	Q9ENP6 hepatitis a
38	95	94.1	251	12	Q9ENQ8	Q9ENQ8 hepatitis a
39	95	94.1	251	12	Q9ENP3	Q9ENP3 hepatitis a
40	95	94.1	251	12	Q9ENN8	Q9ENN8 hepatitis a
41	95	94.1	251	12	Q9ENQ0	Q9ENQ0 hepatitis a
42	95	94.1	2218	12	Q67817	Q67817 hepatitis a
43	95	94.1	2227	12	Q8QV03	Q8QV03 hepatitis a
44	94	93.1	184	12	Q87092	Q87092 simian hepa
45	94	93.1	2225	12	Q9DWK1	Q9DWK1 hepatitis a

ALIGNMENTS

RESULT 1  
ID Q9ENQ4 PRELIMINARY; PRT: 251 AA.  
AC Q9ENQ4;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A201;  
RA Fujiwara K.;  
RT "Hepatitis A virus."  
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047659; BAB12167.1; .  
FT NON\_TER 1  
FT NON\_TER 251  
SQ SEQUENCE 251 AA; 28720 MW; C3342482862F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 1e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEETANSKDFPNMSETDL 20  
Db 95 DLEETANSKDFPNMSETDL 114

RESULT 2

ID Q9ENN2 PRELIMINARY; PRT: 251 AA.  
AC Q9ENN2;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE Polypotein (Fragment).  
OC Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A9;  
RA Fujiwara K.;  
RT "Hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047681; BAB12189.1; -.  
FT NON\_TER 1  
FT NON\_TER 251  
SQ SEQUENCE 251 AA; 28720 MW; C334248282F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 1e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEBIANSKDFPMSETDL 20  
|||||  
95 DLEBIANSKDFPMSETDL 114

## RESULT 3

Q9ENP2 PRELIMINARY; PRT; 251 AA.  
AC Q9ENP2;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A5;  
RA Fujiwara K.;  
RT "Hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047671; BAB12179.1; -.  
FT NON\_TER 1  
FT NON\_TER 251  
SQ SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 1e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEBIANSKDFPMSETDL 20  
|||||  
95 DLEBIANSKDFPMSETDL 114

## RESULT 4

Q9ENP6 PRELIMINARY; PRT; 251 AA.  
AC Q9ENP6;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A162;  
RA Fujiwara K.;

RT "Hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047657; BAB12165.1; -.  
FT NON\_TER 1  
FT NON\_TER 251  
SQ SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6FD3B CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 1e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEBIANSKDFPMSETDL 20  
|||||  
95 DLEBIANSKDFPMSETDL 114

## RESULT 5

Q9ENR1 PRELIMINARY; PRT; 251 AA.  
AC Q9ENR1;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A1;  
RA Fujiwara K.;  
RT "Hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047652; BAB12160.1; -.  
FT NON\_TER 1  
FT NON\_TER 251  
SQ SEQUENCE 251 AA; 28749 MW; 58A520D873893445 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 1e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEBIANSKDFPMSETDL 20  
|||||  
95 DLEBIANSKDFPMSETDL 114

## RESULT 6

Q9ENP1 PRELIMINARY; PRT; 251 AA.  
AC Q9ENP1;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A503;  
RA Fujiwara K.;  
RT "Hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047672; BAB12180.1; -.  
FT NON\_TER 1  
FT NON\_TER 251  
SQ SEQUENCE 251 AA; 28663 MW; C7EA66FBD19A1619 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 1e-07;



Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEEIAANSKDFPNMSETDL 20  
|||||  
95 DLEEIAANSKDFPNMSETDL 114

Db

RESULT 7

O9ENQ1 PRELIMINARY; PRT; 251 AA.

AC O9ENQ1: 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
[1]  
SEQUENCE FROM N.A.  
RA Fujikawa K.;  
RT "hepatitis A virus."  
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047662; BAB12170.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 251 AA; 28699 MW; 8EADAE7E2754C37 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 1e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEEIAANSKDFPNMSETDL 20  
|||||  
95 DLEEIAANSKDFPNMSETDL 114

Db

RESULT 8

O9ENQ9 PRELIMINARY; PRT; 251 AA.

AC O9ENQ9: 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
[1]  
SEQUENCE FROM N.A.  
RA Fujikawa K.;  
RT "hepatitis A virus."  
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047654; BAB12162.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 1e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEEIAANSKDFPNMSETDL 20  
|||||  
95 DLEEIAANSKDFPNMSETDL 114

Db

RESULT 9

O9ENNA

ID O9ENNA PRELIMINARY; PRT; 251 AA.

AC O9ENNA: 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
[1]  
SEQUENCE FROM N.A.  
RA Fujikawa K.;  
RT "hepatitis A virus."  
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047679; BAB12187.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 251 AA; 28734 MW; 3B96B782882F19D9 CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 1e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEEIAANSKDFPNMSETDL 20  
|||||  
95 DLEEIAANSKDFPNMSETDL 114

Db

RESULT 10

O9ENP5 PRELIMINARY; PRT; 251 AA.

AC O9ENP5: 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
[1]  
SEQUENCE FROM N.A.  
RA Fujikawa K.;  
RT "hepatitis A virus."  
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047668; BAB12176.1; -.  
FT NON\_TER 1 1  
SQ SEQUENCE 251 AA; 28614 MW; 8334EF179C757A6D CRC64;

Query Match 100.0%; Score 101; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 1e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DLEEIAANSKDFPNMSETDL 20  
|||||  
95 DLEEIAANSKDFPNMSETDL 114

Db

RESULT 11

O9ENP7 PRELIMINARY; PRT; 251 AA.

AC O9ENP7: 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.

```
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A306;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047666; BAB12174.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||
95 DLEEIAANSKDFPMSETDL 114

SOUT 12
Q9ENN5 PRELIMINARY; PRT; 251 AA.
AC Q9ENN5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A20;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047658; BAB12166.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||
95 DLEEIAANSKDFPMSETDL 114

RESULT 13
Q9ENP9 PRELIMINARY; PRT; 251 AA.
AC Q9ENP9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A303;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047664; BAB12172.1; -.
FT NON_TER 1
FT NON_TER 1
```

```
FT NON_TER 251
SQ SEQUENCE 251 AA; 28752 MW; 7215A28AD2CA5C1A CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||
95 DLEEIAANSKDFPMSETDL 114

RESULT 14
Q9ENN6 PRELIMINARY; PRT; 251 AA.
AC Q9ENN6;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A713;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047677; BAB12185.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||
95 DLEEIAANSKDFPMSETDL 114

RESULT 15
Q9ENN5 PRELIMINARY; PRT; 251 AA.
AC Q9ENN5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A75;
RA Fujiwara K.;
RT "hepatitis A virus.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047678; BAB12186.1; -.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 251 AA; 28658 MW; 98E8BEDD00B2EDF10 CRC64;

Query Match
Best Local Similarity 100.0%; Score 101; DB 12; Length 251;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 DLEEIAANSKDFPMSETDL 20
|||||
```

Thu Oct 2 09:14:00 2003

us-09-171-432a-45.rspt

Page 5

Db 95 DLEBIANSKDPNNSETDL 114

Search completed: October 1, 2003, 10:02:40  
Job time : 33.1765 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 : Search time 12.7059 Seconds  
(without alignments)  
66,600 Million cell updates/sec

Title: US-09-171-432a-45

Perfect score: 101

Sequence: 1 DLEETANSKDFPNMSETDL 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Minimum number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCRTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	2227	3	US-08-475-886-2
2	101	100.0	2227	3	US-08-475-886-4
3	101	100.0	2227	3	US-08-475-886-6
4	101	100.0	2227	3	US-08-397-232-2
5	101	100.0	2227	3	US-08-397-232-4
6	101	100.0	2227	3	US-09-171-387-2
7	101	100.0	2227	3	US-09-653-499-2
8	101	100.0	2227	4	US-09-653-499-4
9	101	100.0	2227	4	US-09-653-499-6
10	44	43.6	907	3	US-08-938-830-26
11	44	43.6	907	3	US-09-020-222-26
12	43	42.6	1201	3	US-09-098-901-2
13	42	41.6	365	2	US-08-983-352-2
14	42	41.6	416	4	US-09-107-532A-5551
15	41	40.6	376	3	US-09-200-965-2
16	41	40.6	508	4	US-09-252-991A-18910
17	40.5	40.1	191	4	US-09-356-806-8
18	40.5	39.6	496	4	US-09-198-452A-1095
19	40.5	39.6	496	4	US-09-339-159B-281
20	39.5	39.1	224	4	US-09-134-001C-3648
21	39	38.6	50	1	US-08-127-351-14
22	39	38.6	50	1	US-08-480-367B-14
23	39	38.6	50	1	US-08-487-221A-14
24	39	38.6	50	1	US-08-480-370-14
25	39	38.6	135	4	US-09-198-452A-544
26	39	38.6	168	4	US-09-134-001C-4664
27	39	38.6	288	4	US-09-107-532A-6783

28	39	38.6	352	3	US-09-286-691-26	Sequence 26, Appl
29	39	38.6	352	3	US-09-687-147-26	Sequence 26, Appl
30	39	38.6	387	4	US-09-314-847A-4	Sequence 4, Appl1
31	39	38.6	387	4	US-09-570-778A-2	Sequence 2, Appl1
32	39	38.6	387	4	US-09-570-778A-7	Sequence 7, Appl1
33	39	38.6	387	4	US-09-570-778A-8	Sequence 8, Appl1
34	39	38.6	387	4	US-09-570-778A-9	Sequence 9, Appl1
35	39	38.6	387	4	US-09-591-138-2	Sequence 2, Appl1
36	39	38.6	387	4	US-09-591-138-7	Sequence 7, Appl1
37	39	38.6	387	4	US-09-591-138-8	Sequence 8, Appl1
38	39	38.6	387	4	US-09-591-138-9	Sequence 9, Appl1
39	39	38.6	391	4	US-09-482-273-151	Sequence 151, App
40	39	38.6	396	1	US-08-430-024-2	Sequence 2, Appl1
41	39	38.6	396	1	US-08-782-009-2	Sequence 2, Appl1
42	39	38.6	396	3	US-09-017-302-2	Sequence 2, Appl1
43	39	38.6	414	4	US-09-482-273-239	Sequence 239, App
44	39	38.6	984	1	US-08-242-932-2	Sequence 2, Appl1
45	39	38.6	984	1	US-08-714-481-2	Sequence 2, Appl1

#### ALIGNMENTS

```
RESULT 1
US-08-475-886-2
Sequence 2, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2
Query Match
Best Local Similarity 100.0%; Score 101; DB 3; Length 2227;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DLEETANSKDFPNMSETDL 20
DB 931 DLEETANSKDFPNMSETDL 950
RESULT 2
US-08-475-886-4
Sequence 4, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
```

```
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated HAV (pass 35), strain HM-175
US-08-475-886-4
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 DLEIANSKDFPMSETDL 20
Db      931 DLEIANSKDFPMSETDL 950
```

## RESULT 3

```
US-08-475-886-6
Sequence 6, Application US/08475886A
Patent No. 6113912
```

```
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 DLEIANSKDFPMSETDL 20
Db      931 DLEIANSKDFPMSETDL 950
```

## RESULT 4

```
US-08-397-232-2
Sequence 2, Application US/08397232A
Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; EARLIER FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
```

```
; TYPE: PRF
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-397-232-2
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 DLEIANSKDFPMSETDL 20
Db      931 DLEIANSKDFPMSETDL 950
```

## RESULT 5

```
US-08-397-232-4
Sequence 4, Application US/08397232A
Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; EARLIER FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-397-232-4
```

```
Query Match          100.0%; Score 101; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 DLEIANSKDFPMSETDL 20
Db      931 DLEIANSKDFPMSETDL 950
```

## RESULT 6

```
US-09-171-387-2
Sequence 2, Application US/09171387
Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; APPLICANT: EMERSON, SUZANNE, U.;
; APPLICANT: PURCELL, ROBERT, H.;
; TITLE OF INVENTION: SIMIAN-HUMAN HAV
; HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US60/015,642
; FILING DATE: 19-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: William S. Feller
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-42290S1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 2
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2227 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
-09-171-387-2
```

```

Query Match          100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3,5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 DLEETIANSKDFPNMSETDL 20
DB      931 DLEETIANSKDFPNMSETDL 950
```

```

RESULT 7
US-09-653-499-2
; Sequence 2, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PORCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2
```

```

Query Match          100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3,5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 DLEETIANSKDFPNMSETDL 20
DB      931 DLEETIANSKDFPNMSETDL 950
```

```

RESULT 8
US-09-653-499-4
; Sequence 4, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
```

```

; APPLICANT: PORCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4
```

```

Query Match          100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3,5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 DLEETIANSKDFPNMSETDL 20
DB      931 DLEETIANSKDFPNMSETDL 950
```

```

RESULT 9
US-09-653-499-6
; Sequence 6, Application US/09653499
; Patent No. 6423318
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PORCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/09/653,499
; CURRENT FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: 08/475,886
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6
```

```

Query Match          100.0%; Score 101; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 3,5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 DLEETIANSKDFPNMSETDL 20
DB      931 DLEETIANSKDFPNMSETDL 950
```

```

RESULT 10
US-08-938-830-26
; Sequence 26, Application US/08938830
; Patent No. 6040437
; GENERAL INFORMATION:
; APPLICANT: Lasky, Laurence A.
; APPLICANT: Dowbenko, Donald J.
; TITLE OF INVENTION: Tyrosine Phosphorylated Cleavage
```





COUNTRY: USA  
ZIP: 22201-4741  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/983,352  
FILING DATE: 02-FEB-1998  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/GB96/01629  
FILING DATE: 08-JUL-1996  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: GB 9514138.8  
FILING DATE: 11-JUL-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: GB 9607220.2  
FILING DATE: 04-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Crawford, Arthur R.  
REGISTRATION NUMBER: 25,327  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-816-4000  
TELEFAX: 703-816-4100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 365 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-983-352-2

Query Match 41.6%; Score 42; DB 2; Length 365;  
Best Local Similarity 47.4%; Pred. No. 33;  
Matches 9; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 2 LEEIANSKDPNMSSTD 20  
DB 258 IEELAKRGIAVLMSTDL 276

RESULT 14  
US-09-107-532A-5551  
Sequence 5551, Application US/09107532A  
Patent No. 6583275  
GENERAL INFORMATION:  
APPLICANT: Lynn A Doucelte-Stamm and David Bush  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
NUMBER OF SEQUENCES: 7310  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GENEOME THERAPEUTICS CORPORATION  
STREET: 100 Beaver Street  
CITY: Waltham  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02354  
COMPUTER READABLE FORM:  
MEDIUM TYPE: CD-ROM ISO9660  
COMPUTER: PC  
OPERATING SYSTEM: <Unknown>  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,532A  
FILING DATE: 30-Jun-1998  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 60/085,598  
FILING DATE: 14 May 1998  
APPLICATION NUMBER: 60/051571  
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:  
NAME: Arianello, Pamela Deneke  
REGISTRATION NUMBER: 40,489  
REFERENCE/DOCKET NUMBER: GTC-012  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781)893-5007  
TELEFAX: (781)893-8277  
INFORMATION FOR SEQ ID NO: 5551:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 416 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: YES  
ORIGINAL SOURCE:  
ORGANISM: Enterococcus faecium  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: (B) LOCATION 1...416  
SEQUENCE DESCRIPTION: SEQ ID NO: 5551:  
US-09-107-532A-5551

Query Match 41.6%; Score 42; DB 4; Length 416;  
Best Local Similarity 53.8%; Pred. No. 39;  
Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 3 LEEIANSKDPNPM 15  
DB 117 ERYLTNTKDPNL 129

RESULT 15  
US-09-200-965-2  
Sequence 2, Application US/09200965  
Patent No. 6134422  
GENERAL INFORMATION:  
APPLICANT: ROSEN, ET AL.  
TITLE OF INVENTION: Thrombin Inhibitor  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/200,965  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/171,817  
FILING DATE: 22-DECEMBER-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-47  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN

Thu Oct 2 09:13:59 2003

us-09-171-432a-45.raii

Page 6

US-09-200-965-2

Query Match	40.6%	Score 41	DB 3	Length 376
Best Local Similarity	33.3%	Pred, No. 50		
Matches 10, Conservative	3	Mismatches 7	Indels 10	Gaps 1

```
QY      1 DLEEIAAN-----SKDFPNMSETDL 20
        1:1: | | | | |
DB      281 DMESVLRNLGNTDAFELGKADFGSMQTDL 310
```

Search completed: October 1, 2003, 10:06:32  
Job time : 13.7059 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 ; Search time 24.1176 Seconds  
(without alignments)  
131.201 Million cell updates/sec

Title: US-09-171-432a-45

Perfect score: 101

Sequence: 1 DLEIRIANSKDFPNMSETDL 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PC1\_NEM\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEM\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEM\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PC1US\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEM\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEM\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEM\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEM\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	100.0	352	US-10-272-459-45	Sequence 45, Appl
2	101	100.0	980	US-10-272-459-41	Sequence 41, Appl
3	101	100.0	2227	US-09-929-955-12	Sequence 12, Appl
4	101	100.0	2227	US-10-104-966-12	Sequence 12, Appl
5	101	100.0	2227	US-10-135-988-2	Sequence 2, Appl
6	101	100.0	2227	US-10-135-988-4	Sequence 6, Appl
7	101	100.0	2227	US-10-135-988-6	Sequence 27, Appl
8	44	43.6	401	US-10-195-144-27	Sequence 34, Appl
9	42	41.6	538	US-09-769-744A-34	Sequence 41631, A
10	41	40.6	72	US-09-864-761-41831	Sequence 1440, Ap
11	41	40.6	459	US-09-925-300-1440	Sequence 8, Appl
12	40	39.6	528	US-10-205-522-8	Sequence 1, Appl
13	40	39.6	284	US-09-810-997-1	Sequence 1, Appl
14	40	39.6	284	US-10-174-209-1	Sequence 11280, A
15	40	39.6	381	US-09-815-242-11280	

16	40	39.6	499	9	US-09-864-761-35385	Sequence 35385, A
17	40	39.6	569	15	US-10-156-761-12273	Sequence 12273, A
18	40	39.6	715	12	US-10-032-585-7002	Sequence 7002, Ap
19	40	39.6	786	12	US-10-263-568-8	Sequence 8, Appl
20	40	39.6	1230	12	US-10-205-219-44	Sequence 44, Appl
21	40	39.6	1230	12	US-10-205-219-97	Sequence 97, Appl
22	39	38.6	337	12	US-10-174-209-6	Sequence 6, Appl
23	39	38.6	331	12	US-09-803-286A-10	Sequence 10, Appl
24	39	38.6	341	9	US-10-166-225A-87	Sequence 87, Appl
25	39	38.6	341	9	US-09-815-242-10657	Sequence 10657, A
26	39	38.6	387	11	US-09-991-138-2	Sequence 2, Appl
27	39	38.6	387	11	US-09-991-138-7	Sequence 7, Appl
28	39	38.6	387	11	US-09-991-138-8	Sequence 8, Appl
29	39	38.6	387	11	US-09-991-138-9	Sequence 9, Appl
30	39	38.6	387	11	US-10-037-677-4	Sequence 4, Appl
31	39	38.6	391	11	US-09-984-271-151	Sequence 151, App
32	39	38.6	396	9	US-09-815-242-10387	Sequence 10387, A
33	39	38.6	414	11	US-09-984-271-239	Sequence 239, App
34	39	38.6	417	9	US-09-815-242-10492	Sequence 10492, A
35	39	38.6	420	10	US-09-764-868-912	Sequence 912, App
36	39	38.6	569	15	US-10-034-934-87	Sequence 87, Appl
37	39	38.6	722	15	US-10-222-668-4	Sequence 4, Appl
38	39	38.6	978	10	US-09-893-817-4	Sequence 4, Appl
39	39	38.6	978	10	US-09-893-817-8	Sequence 8, Appl
40	39	38.6	1111	9	US-09-815-242-12955	Sequence 12955, A
41	39	38.6	1247	9	US-09-803-286A-2	Sequence 2, Appl
42	38.5	38.1	383	12	US-10-037-585-7513	Sequence 7513, App
43	38	37.6	89	12	US-10-440-066-36	Sequence 36, Appl
44	38	37.6	89	12	US-10-440-066-37	Sequence 37, Appl
45	38	37.6	89	12	US-10-440-066-38	Sequence 38, Appl

## ALIGNMENTS

RESULT 1  
US-10-272-459-45  
; Sequence 45, Application US/10272459  
; Publication No. US20030124517A1  
; GENERAL INFORMATION:  
; APPLICANT: PICHUANES, Sergio  
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT  
; FILE REFERENCE: PPI7955.002 / 2301-17955  
; CURRENT APPLICATION NUMBER: US/10/272,459  
; CURRENT FILING DATE: 2002-10-15  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein  
; OTHER INFORMATION: of 38.8 kDa  
US-10-272-459-45

Query Match 100.0%: Score 101; DB 15; Length 352;  
Best Local Similarity 100.0%: Pred. No. 1.6e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 DLEIRIANSKDFPNMSETDL 20  
DB 303 DLEIRIANSKDFPNMSETDL 322  
US-10-272-459-41  
; Sequence 41, Application US/10272459  
; Publication No. US20030124517A1  
; GENERAL INFORMATION:  
; APPLICANT: PICHUANES, Sergio  
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

```
;; TITLE OF INVENTION: PROTEINS AND USES THEREOF
;; FILE REFERENCE: PP17955.002 / 2301-17955
;; CURRENT APPLICATION NUMBER: US/10/272,459
;; CURRENT FILING DATE: 2002-10-15
;; NUMBER OF SEQ ID NOS: 50
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 41
;; LENGTH: 980
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41
```

```
Query Match          100.0%; Score 101; DB 15; Length 980;
Best Local Similarity 100.0%; Pred. No. 5.3e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 DLEIANSKDFPMSETDL 20
        |||
DB      931 DLEIANSKDFPMSETDL 950
```

```
RESULT 3
US-09-929-955-12
;; Sequence 12, Application US/09929955
;; Patent No. US20020136740A1
;; GENERAL INFORMATION:
;; APPLICANT: Matti Salberg
;; APPLICANT: Catharina Hultgren
;; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
;; TITLE OF INVENTION: METHODS OF USE THEREOF
;; FILE REFERENCE: TRIPER.23AUS2
;; CURRENT APPLICATION NUMBER: US/09/929,955
;; CURRENT FILING DATE: 2001-08-15
;; PRIOR APPLICATION NUMBER: 09/705,547
;; PRIOR FILING DATE: 2000-11-03
;; PRIOR APPLICATION NUMBER: 60/229,175
;; PRIOR FILING DATE: 2000-08-29
;; PRIOR APPLICATION NUMBER: 60/225,767
;; PRIOR FILING DATE: 2000-08-17
;; NUMBER OF SEQ ID NOS: 49
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 12
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Hepatitis A virus sequence
09-929-955-12
```

```
Query Match          100.0%; Score 101; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 DLEIANSKDFPMSETDL 20
        |||
DB      931 DLEIANSKDFPMSETDL 950
```

```
RESULT 4
US-10-104-966-12
;; Sequence 12, Application US/10104966
;; Publication No. US2002015124A1
;; GENERAL INFORMATION:
;; APPLICANT: Matti Salberg
;; APPLICANT: Catharina Hultgren
;; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
;; TITLE OF INVENTION: METHODS OF USE THEREOF
;; FILE REFERENCE: TRIPER.23AUS1
;; CURRENT APPLICATION NUMBER: US/10/104,966
;; CURRENT FILING DATE: 2002-03-22
```

```
;; PRIOR APPLICATION NUMBER: 09/705,547
;; PRIOR FILING DATE: 2000-11-03
;; PRIOR APPLICATION NUMBER: 60/229,175
;; PRIOR FILING DATE: 2000-08-29
;; NUMBER OF SEQ ID NOS: 15
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 12
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12
```

```
Query Match          100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 DLEIANSKDFPMSETDL 20
        |||
DB      931 DLEIANSKDFPMSETDL 950
```

```
RESULT 5
US-10-135-988-2
;; Sequence 2, Application US/10135988
;; Publication No. US20020176869A1
;; GENERAL INFORMATION:
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US3
;; CURRENT APPLICATION NUMBER: US/10/135,988
;; CURRENT FILING DATE: 2002-04-29
;; PRIOR APPLICATION NUMBER: 07/947,338
;; PRIOR FILING DATE: 1992-09-18
;; PRIOR APPLICATION NUMBER: 08/397,232
;; PRIOR FILING DATE: 1995-03-10
;; NUMBER OF SEQ ID NOS: 6
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 2
;; LENGTH: 2227
;; TYPE: PRT
;; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HN-175
US-10-135-988-2
```

```
Query Match          100.0%; Score 101; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 DLEIANSKDFPMSETDL 20
        |||
DB      931 DLEIANSKDFPMSETDL 950
```

```
RESULT 6
US-10-135-988-4
;; Sequence 4, Application US/10135988
;; Publication No. US20020176869A1
;; GENERAL INFORMATION:
;; APPLICANT: FUNKHOUSER, ANN W
;; APPLICANT: EMERSON, SUZANNE U
;; APPLICANT: PURCELL, ROBERT H
;; APPLICANT: D'HONDT, ERIC
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
;; FILE REFERENCE: 20264262US3
;; CURRENT APPLICATION NUMBER: US/10/135,988
;; CURRENT FILING DATE: 2002-04-29
;; PRIOR APPLICATION NUMBER: 07/947,338
;; PRIOR FILING DATE: 1992-09-18
;; PRIOR APPLICATION NUMBER: 08/397,232
```

;; PRIOR FILING DATE: 1995-03-10  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 4  
;; LENGTH: 2227  
;; TYPE: PRF  
;; ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
US-10-135-988-4

Query Match 100.0%; Score 101; DB 14;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETANSKDFPNMSETDL 20  
DB 931 DLEETANSKDFPNMSETDL 950

RESULT 7  
US-10-135-988-6  
Sequence 6, Application US/10135988  
Publication No. US20020176669A1

;; GENERAL INFORMATION:  
;; APPLICANT: FUNKHOUSER, ANN W  
;; APPLICANT: EMERSON, SOZANNE U  
;; APPLICANT: PURCELL, ROBERT H  
;; APPLICANT: D'HONDT, ERIC  
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
;; FILE REFERENCE: 20264262053  
;; CURRENT APPLICATION NUMBER: US/10/135,988  
;; CURRENT FILING DATE: 2002-04-29  
;; PRIOR APPLICATION NUMBER: 07/947,338  
;; PRIOR FILING DATE: 1992-09-18  
;; PRIOR APPLICATION NUMBER: 08/397,232  
;; PRIOR FILING DATE: 1995-03-10  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 6  
;; LENGTH: 2227  
;; TYPE: PRF  
;; ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-10-135-988-6

Query Match 100.0%; Score 101; DB 14;  
Best Local Similarity 100.0%; Pred. No. 1.4e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLEETANSKDFPNMSETDL 20  
DB 931 DLEETANSKDFPNMSETDL 950

RESULT 8  
US-10-195-144-27  
Sequence 27, Application US/10195144  
Publication No. US2003012666A1  
;; GENERAL INFORMATION:  
;; APPLICANT: BROWN, GREGORY G.  
;; APPLICANT: FORMANOVA, NATASA  
;; APPLICANT: DENDY, CHARLES  
;; APPLICANT: LANDRY, BENJOIT S.  
;; APPLICANT: CHEUNG, WING  
;; APPLICANT: JIN, HUA  
;; TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN  
;; TITLE OF INVENTION: PLANTS  
;; FILE REFERENCE: 16313-0136  
;; CURRENT APPLICATION NUMBER: US/10/195,144  
;; CURRENT FILING DATE: 2002-10-01  
;; PRIOR APPLICATION NUMBER: 60/305,026  
;; PRIOR FILING DATE: 2001-07-12  
;; PRIOR APPLICATION NUMBER: 60/305,363  
;; PRIOR FILING DATE: 2001-07-13  
;; PRIOR APPLICATION NUMBER: 60/308,736

;; PRIOR FILING DATE: 2001-07-30  
;; NUMBER OF SEQ ID NOS: 128  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 27  
;; LENGTH: 401  
;; TYPE: PRF  
;; ORGANISM: Raphanus sativum  
US-10-195-144-27

Query Match 43.6%; Score 44; DB 15; Length 401;  
Best Local Similarity 56.2%; Pred. No. 44;  
Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 DLEETANSKDFPNM 16  
DB 9 ELDEETALSKEPMSLS 24

RESULT 9  
US-09-769-744A-34  
Sequence 34, Application US/09769744A  
Publication No. US20030134407A1

;; GENERAL INFORMATION:  
;; APPLICANT: Le Page, Richard WF  
;; APPLICANT: Wells, Jeremy M  
;; APPLICANT: Hanniffy, Sean B  
;; APPLICANT: Hansbro, Philip M  
;; TITLE OF INVENTION: Proteins  
;; FILE REFERENCE: PWC/P21122WO  
;; CURRENT APPLICATION NUMBER: US/09/769,744A  
;; CURRENT FILING DATE: 2001-01-26  
;; PRIOR APPLICATION NUMBER: PCT/GB99/02452  
;; PRIOR FILING DATE: 1999-07-27  
;; PRIOR APPLICATION NUMBER: GB 9816336.3  
;; PRIOR FILING DATE: 1998-07-27  
;; PRIOR APPLICATION NUMBER: US 60/125329  
;; PRIOR FILING DATE: 1999-03-19  
;; NUMBER OF SEQ ID NOS: 196  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 34  
;; LENGTH: 538  
;; TYPE: PRF  
;; ORGANISM: Streptococcus pneumoniae  
US-09-769-744A-34

Query Match 41.6%; Score 42; DB 12; Length 538;  
Best Local Similarity 53.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 LEEETANSKDFPN 14  
DB 436 LEOVGSLVKDFPN 448

RESULT 10  
US-09-864-761-41831  
Sequence 41831, Application US/09864761  
Patent No. US20020046763A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Penn, Sharon G.  
;; APPLICANT: Rank, David R.  
;; APPLICANT: Hanzel, David K.  
;; APPLICANT: Chen, Wensheng  
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
;; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
;; FILE REFERENCE: Aecomica-X-1  
;; CURRENT APPLICATION NUMBER: US/09/864,761  
;; CURRENT FILING DATE: 2001-05-23  
;; PRIOR APPLICATION NUMBER: US 60/180,312  
;; PRIOR FILING DATE: 2000-02-04  
;; PRIOR APPLICATION NUMBER: US 60/207,456  
;; PRIOR FILING DATE: 2000-05-26  
;; PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 41831  
LENGTH: 72  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC025488.1  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.2  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3  
OTHER INFORMATION: EST\_HUMAN HIT: BF689099.1, EVALUE 7.00e-34  
OTHER INFORMATION: SWISSPROT HIT: P06133, EVALUE 1.00e-37  
US-09-864-761-41831

Query Match 40.6%; Score 41; DB 9; Length 72;  
Best Local Similarity 66.7%; Pred. No. 18;  
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 6 AANSKDFPNMSETDL 20  
Db 47 AAVSLDFHTMSSTD 61

RESULT 11  
US-09-925-300-1440  
Sequence 1440, Application US/09925300  
Patent No. US20020151681A1  
GENERAL INFORMATION:  
APPLICANT: Craig Rosen,  
APPLICANT: Steve Ruben  
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
FILE REFERENCE: PA101  
CURRENT APPLICATION NUMBER: US/09/925,300  
PRIOR FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: PCT/US00/05988  
PRIOR FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12  
NUMBER OF SEQ ID NOS: 1890  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1440  
LENGTH: 459  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-925-300-1440  
Query Match 40.6%; Score 41; DB 10; Length 459;  
Best Local Similarity 33.3%; Pred. No. 1.6e+02;  
Matches 10; Conservative 3; Mismatches 7; Indels 10; Gaps 1;  
OY 1 DLEBIAAN-----SKDFPNMSETDL 20  
Db 364 DMESVLRNLGWTDAFELGKADFGSMQSTD 393  
RESULT 12  
US-10-205-522-8  
Sequence 8, Application US/10205522  
Publication No. US2003007629A1  
GENERAL INFORMATION:  
APPLICANT: Penny, Laura  
APPLICANT: Galvin, Margaret  
APPLICANT: Miller, Andrew  
APPLICANT: Reidy, Michael  
TITLE OF INVENTION: Genotyping Human  
TITLE OF INVENTION: UDP-Glucuronosyltransferase 2B4 (UGT2B4), 2B7 (UGT2B7) and  
FILE REFERENCE: SEQ-22PRV2  
CURRENT APPLICATION NUMBER: US/10/205,522  
PRIOR FILING DATE: 2002-07-24  
PRIOR APPLICATION NUMBER: US/09/356,806  
PRIOR FILING DATE: 1999-07-20  
NUMBER OF SEQ ID NOS: 164  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 8  
LENGTH: 528  
TYPE: PRT  
ORGANISM: H. sapiens  
US-10-205-522-8

Query Match 40.6%; Score 41; DB 15; Length 528;  
Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 6 AANSKDFPNMSETDL 20  
Db 411 AAVSLDFHTMSSTD 425

RESULT 13  
US-09-810-997-1  
Sequence 1, Application US/09810997  
Patent No. US20020007501A1  
GENERAL INFORMATION:  
APPLICANT: Song, Xiaoling  
APPLICANT: Fan, Hao  
APPLICANT: Wei, Zhong-Min  
TITLE OF INVENTION: RECEPTORS FOR HYPERSENSITIVE RESPONSE ELICITORS AND  
FILE REFERENCE: 21829/62  
CURRENT APPLICATION NUMBER: US/09/810,997  
PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 60/191,649  
PRIOR FILING DATE: 2000-03-23  
PRIOR APPLICATION NUMBER: 60/250,710  
PRIOR FILING DATE: 2000-12-01  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 284

TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-810-997-1

Query Match 39.6%; Score 40; DB 9; Length 284;  
Best Local Similarity 44.4%; Pred. No. 1.3e+02;  
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 3 EEIANSKDFPNMSETDL 20  
DB 174 GRIDVFSKDFDNIAEVEL 191

RESULT 14  
US-10-174-209-1  
Sequence 1, Application US/10174209  
Publication No. US20030177526A1

GENERAL INFORMATION:  
APPLICANT: Song, Xiaoling

APPLICANT: Bartola, Pauline A.

APPLICANT: Linderoch, No. US20030177526A1 A.

APPLICANT: Fan, Hao

APPLICANT: Wei, Zhong-Min

TITLE OF INVENTION: RECEPTORS FOR HYPERSENSITIVE RESPONSE ELICITORS AND

TITLE OF INVENTION: USES THEREOF

FILE REFERENCE: 21829/211

CURRENT APPLICATION NUMBER: US/10/174,209

CURRENT FILING DATE: 2002-06-17

PRIOR APPLICATION NUMBER: 60/335,776

PRIOR FILING DATE: 2001-10-31

PRIOR APPLICATION NUMBER: 09/810,997

PRIOR FILING DATE: 2001-03-16

NUMBER OF SEQ ID NOS: 86

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 1

LENGTH: 284

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-10-174-209-1

Query Match 39.6%; Score 40; DB 12; Length 284;  
Best Local Similarity 44.4%; Pred. No. 1.3e+02;  
Matches 8; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

OY 3 EEIANSKDFPNMSETDL 20

DB 174 GRIDVFSKDFDNIAEVEL 191

RESULT 15  
US-09-815-242-11280  
Sequence 11280, Application US/09815242  
Patent No. US20020061569A1

GENERAL INFORMATION:  
APPLICANT: Haselbeck, Robert

APPLICANT: Ohlsen, Karl L.

APPLICANT: Zyskind, Judith W.

APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.

APPLICANT: Carr, Grant J.

APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard

TITLE OF INVENTION: Identification of Essential Genes in

TITLE OF INVENTION: Prokaryotes

FILE REFERENCE: ELITRA.011A

CURRENT APPLICATION NUMBER: US/09/815,242

CURRENT FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 11280

LENGTH: 381

TYPE: PRT

ORGANISM: Haemophilus influenzae

US-09-815-242-11280

Query Match 39.6%; Score 40; DB 9; Length 381;  
Best Local Similarity 47.4%; Pred. No. 1.9e+02;  
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 1 DLEIANSKDFPNMSETD 19

DB 41 DLEIANSKDFPNMSETD 59

Search completed: October 1, 2003, 10:37:52

Job time : 24.1176 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 ; Search time 43.5294 Seconds  
(without alignments)  
72.928 Million cell updates/sec

Title: US-09-171-432a-46

Perfect score: 96  
Sequence: 1 KINLADRMGLSGVQEKREQ 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database :  
1: A.Geneseq.19Jun03.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	20	AAW42929	Immunogenic Hepat
2	96	100.0	20	AAW42929	Synthetic HAV P2A
3	96	100.0	25	AAW42969	Immunogenic Hepat
4	96	100.0	25	AAW42969	Synthetic HAV P2A
5	96	100.0	2227	AAW42969	Sequence of viral
6	96	100.0	2227	AAW42969	Attenuated hepati
7	96	100.0	2227	AAW42969	Hepatitis A virus
8	96	100.0	2227	AAW42969	Amino acid sequenc
9	96	100.0	2227	AAW42969	Amino acid sequenc

10	96	100.0	2227	21	AAW42929	standard; peptide; 20 AA.
11	96	100.0	2227	23	AAW42929	(first entry)
12	96	100.0	2227	23	AAW42929	Immunogenic Hepatitis A virus peptide YK-1331.
13	96	100.0	2227	23	AAW42929	Immunogenic peptide; immunogenic epitope; P2A protein;
14	96	100.0	2227	23	AAW42929	Immune response; antibody.
15	96	100.0	2227	23	AAW42929	Synthetic.
16	96	100.0	2227	23	AAW42929	Hepatitis A virus.
17	96	100.0	2227	23	AAW42929	WO9740147-A1.
18	96	100.0	2227	23	AAW42929	30-OCT-1997.
19	96	100.0	2227	23	AAW42929	18-APR-1997; 97MO-US06891.
20	96	100.0	2227	23	AAW42929	19-APR-1996; 96US-0015644.
21	96	100.0	2227	23	AAW42929	(USSH) US DEPT HEALTH & HUMAN SERVICES.
22	96	100.0	2227	23	AAW42929	Fields HA, Khudiyakov YE;
23	96	100.0	2227	23	AAW42929	WPI; 1997-535831/49.
24	96	100.0	2227	23	AAW42929	Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an
25	96	100.0	2227	23	AAW42929	Immune response to HAV in a mammal or to detect the presence of

## ALIGNMENTS

Amino acid sequenc  
Wild-type Hepatiti  
Hepatitis A virus  
Attenuated Hepatit  
Hepatitis A virus  
Wild type human he  
Attenuated (pass3  
Attenuated hepati  
Propionibacterium  
Streptococcus poly  
Shrimp white spot  
Streptococcus poly  
Lactococcus lactis  
Propionibacterium  
Saccharomyces cere  
Amino acid sequenc  
Human kinase (PKIN  
Human steroid and  
Human steroid and  
Drosophila melanog  
Novel human diagno  
Protein which is s  
Mouse CD14 protein  
A murine CD14 prot  
Pseudomonas aerugi  
Chlamydia trachoma  
Amino acid sequenc  
Human lung cancer  
Human clone L761P  
Novel human diagno  
Cyanobacterial phy  
Novel human diagno  
C glutamicum prote  
Propionibacterium  
Propionibacterium  
Human liver peptid

PT antibodies against HAV in a mammal  
XX  
XX Claim 18; Page 112; 140pp; English.  
XX  
CC Peptides AAW42922-30 are immunogenic peptides corresponding to  
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are  
CC substantially similar to a portion of the amino acid sequence of the P2A  
CC protein of HAV corresponding to amino acids 961-980. The present peptide  
CC is derived from amino acids 961-980, and has a reactivity of 27.1% with  
CC acute sera. Compositions containing the peptides can be used to induce an  
CC immune response to HAV in a mammal. The peptides can also be used to  
CC detect the presence of antibodies against HAV in mammalian serum. The  
CC peptides can also be used to make an antibody against HAV by  
CC administering the peptide to a mammal.  
XX  
SQ Sequence 20 AA;  
Query Match 100.0%; Score 96; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 9.2e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KINLADRLGLSGVOEIKQ 20  
1 KINLADRLGLSGVOEIKQ 20  
Db  
RESULT 2  
AAB69446  
ID AAB69446 standard; Peptide: 20 AA.  
XX  
AC AAB69446;  
XX  
XX 20-APR-2001 (first entry)  
DT  
XX  
XX Synthetic HAV P2A peptide, SEQ ID NO: 46.  
DE  
XX Hepatitis A virus; HAV, immunogen; immunostimulant; virucide; vaccine;  
KM antigen; major structural capsid polypeptide; HAV antibody detection.  
KW  
XX Hepatitis A virus.  
OS  
XX Synthetic.  
OS  
XX WO200105824-A2.  
PN  
XX 25-JAN-2001.  
PD  
XX 14-JUL-2000; 2000WO-US19267.  
PF  
XX 15-JUL-1999; 99US-0144412.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Fields HA, Khudyakov YE;  
PI  
XX WPI; 2001-112681/12.  
DR  
XX  
XX Synthetic peptides used as antigen sources for enzyme immunoassays  
PT detecting anti-hepatitis A virus and as vaccines -  
XX  
XX Claim 13; Page 97; 130pp; English.  
XX  
XX The present sequence is one of a number of synthetic peptides which are  
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides  
CC comprise antigenic epitopes of the major structural capsid polypeptides  
CC or non-structural polypeptides of HAV with one or more glutamine  
CC molecules at the carboxy end of the peptide. The peptides are used to  
CC detect the presence of antibodies against HAV in mammalian serum, to  
CC detect the presence of HAV in a human or animal through the binding of  
CC the peptide to an antibody, to detect acute phase infection by detecting  
CC Igm antibodies in mammalian serum and detecting convalescence in a  
CC mammal. The peptides are used to detect or quantify HAV antibodies in  
CC samples in clinical or research-based assays using immunoblotting,  
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,

CC tracking of radioactive or bioluminescent markers, chromatography or  
CC electrophoresis. The peptides are used to induce an immune response to  
CC HAV when administered to a human or animal. Glutamine at the carboxy  
CC end of the peptides enhances the Igm antibody reactivity.  
XX  
SQ Sequence 20 AA;  
Query Match 100.0%; Score 96; DB 22; Length 20;  
Best Local Similarity 100.0%; Pred. No. 9.2e-10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KINLADRLGLSGVOEIKQ 20  
1 KINLADRLGLSGVOEIKQ 20  
Db  
RESULT 3  
AAW42969  
ID AAW42969 standard; peptide: 25 AA.  
XX  
AC AAW42969;  
XX  
XX 28-APR-1998 (first entry)  
DT  
XX  
XX Immunogenic Hepatitis A virus peptide YK-1757.  
DE  
XX Immunogenic peptide; immunogenic epitope; P2A protein;  
KW immune response; antibody.  
KW  
XX Synthetic.  
OS  
XX Hepatitis A virus.  
OS  
XX WO9740147-A1.  
PN  
XX 30-OCT-1997.  
PD  
XX 18-APR-1997; 97WO-US06891.  
PF  
XX 19-APR-1996; 96US-0015644.  
PR  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA  
XX  
XX Fields HA, Khudyakov YE;  
PI  
XX WPI; 1997-535831/49.  
DR  
XX  
XX Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an  
PT immune response to HAV in a mammal or to detect the presence of  
PT antibodies against HAV in a mammal  
XX  
XX Claim 18; Page 112; 140pp; English.  
XX  
XX The present immunogenic peptide corresponds to an immunogenic  
CC epitope of the Hepatitis A virus (HAV). The peptide is substantially  
CC similar to a portion of the amino acid sequence of the P2A protein of HAV  
CC corresponding to amino acids 792-980. Compositions containing the  
CC peptide can be used to induce an immune response to HAV in a mammal.  
CC The peptide can also be used to detect the presence of antibodies  
CC against HAV in mammalian serum. The peptide can also be used to make an  
CC antibody against HAV by administering the peptide to a mammal.  
XX  
SQ Sequence 25 AA;  
Query Match 100.0%; Score 96; DB 18; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1.2e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 KINLADRLGLSGVOEIKQ 20  
6 KINLADRLGLSGVOEIKQ 25  
Db  
RESULT 4

```

AA69448
ID   AAB69448 standard; peptide; 25 AA.
AC   AAB69448;
DT     20-APR-2001 (first entry)
DE   Synthetic HAV P2A peptide, SEQ ID NO: 48.
OS   Hepatitis A virus; Immunogen: immunostimulant; virucide; vaccine;
KM   antigen; major structural capsid polypeptide; HAV antibody detection.
XX   Synthetic.
XX   Hepatitis A virus.
XX   WO200105824-A2.
XX   PN
XX   25-JAN-2001.
XX   PD
XX   14-JUL-2000; 2000WO-US19267.
XX   PR   15-JUL-1999; 99US-0144412.
XX   PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX   PI   Fields HA, Khudyakov YE;
XX   DR   WPI: 2001-112681/12.
XX   PT   Synthetic peptides used as antigen sources for enzyme immunoassays
XX   detecting anti-hepatitis A virus and as vaccines .
XX   PS
XX   Claim 13; Page 99; 130pp; English.

The present sequence is one of a number of synthetic peptides which are
immunoreactive with hepatitis A virus (HAV) antibodies. The peptides are
comprise antigenic epitopes of the major structural capsid polypeptides
or non-structural polypeptides of HAV with one or more glutamine
molecules at the carboxy end of the peptide. The peptides are used to
detect the presence of antibodies against HAV in mammalian serum, to
detect the presence of HAV in a human or animal through the binding of
the peptide to an antibody, to detect acute phase infection by detecting
IgM antibodies in mammalian serum and detecting convalescence in a
mammal. The peptides are used to detect or quantify HAV antibodies in
samples in clinical or research-based assays using immunoblotting,
fluorescence in situ hybridisation analysis, gel-mobility shift assays,
tracking of radioactive or bioluminescent markers, chromatography or
electrophoresis. The peptides are used to induce an immune response to
HAV when administered to a human or animal. Glutamine at the carboxy
end of the peptides enhances the IgM antibody reactivity.

SQ   Sequence      25 AA;

Query Match          100.0%; Score 96; DB 22; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches    20; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

QY       1 KINLADRMIGLSGVOEIKEQ 20
         |||||||
DB        6 KINLADRMIGLSGVOEIKEQ 25

RESULT 5
AAP60066
ID   AAP60066 standard; Protein; 2227 AA.
AC   AAP60066;
XX   DT     25-MAR-2003 (updated)
XX   DT     26-JUN-1991 (first entry)
XX   DE   Sequence of viral L434 polypeptide encoded by the complete
XX   nucleotide sequence of the HAV genome.
```

KW		Diagnosis; vaccine; passive immunotherapy.
XX		
OS	Hepatitis A virus.	
XX		
PH	Key	Location/Qualifiers
FT	Region	1..245 /label= P1.1A
FT	Region	246..491
FT	Region	/label= 1B
FT	Region	492..836
FT	Region	/label= 1C
FT	Region	837..980
FT	Region	/label= P2.2A
FT	Region	981..1076
FT	Region	/label= 2B
FT	Region	1077..1422
FT	Region	/label= 2C
FT	Region	1423..1484
FT	Region	/label= P3.3A
FT	Region	1485..1507
FT	Region	/label= 3B
FT	Region	1508..1678
FT	Region	/label= 3C
FT	Region	1679..2227
FT	Region	/label= 3D
XX		
PN	EP199480-A.	
XX		
PD	29-OCT-1986.	
XX		
PP	03-APR-1986;	86EP-0302465.
XX		
PR	03-APR-1985;	85US-0719329.
XX		
PA	(CHIR ) CHIRON CORP.	
PI	Dina D, Potter SJ, Vannest GA, Caput D;	
DR	WP1: 1986-286213/44.	
XX	N-PSDB; AAN60080.	
PT	Hepatitis A virus nucleotide sequence and polypeptide - and use	
PI	in prodn. of vaccines and diagnostic probes	
XX		
PS	Claim 5; Fig 1; 18pp; English.	
CC	AAN60080 and oligonucleotide fragments are useful in detection of	
CC	hepatitis A virus; transformed hosts may be used for expression of	
CC	polypeptides and fragments useful in vaccines without risk of	
CC	infection by the virus or in prodn. of particles which are capable	
CC	of inducing immunocompetent B cells for passive immunotherapy. Pref.	
CC	epitope is derived from AAs 445-657 or 792-848 of the HAV	
CC	polypeptide sequence (AAP60066).	
CC	(Updated on 25-MAR-2003 to correct PA field.)	
XX		
SQ	Sequence	2227 AA;
QY	Query Match	100.0%; Score 96; DB 7; Length 2227;
	Best Local Similarity	100.0%; Pred. No. 2.2e-07;
	Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	1 KINLADRLGSGVGEIREQ 20	
	661 KINLADRLGLSGVGEIREQ 980	
RESULT 6		
AAR05697		
ID	AAR05697 standard: protein: 2227 AA.	
XX		
AC	AAR05697;	
XX		



CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the  
 CC HAV of (3); and (5) a host cell containing the HAV of (3). (1) or  
 CC its RNA transcript, can be used as a vaccine for preventing HAV in  
 CC a mammal. (1) or the RNA transcript can also be used to stimulate  
 CC the production of protective antibodies in the mammal.

XX  
 SQ Sequence 2227 AA;

Query Match 100.0%; Score 96; DB 18; Length 2227;

Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20  
 DB 961 KINLADRMGLSGVOEIKQ 980

# RESULT 8

AAB18607 standard; Protein; 2227 AA.

AC AAB18607;

DT 15-JAN-2001 (first entry)

DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.

KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.

XX Hepatitis A virus.

OS US6113912-A.

PN 05-SEP-2000.

PD 07-JUN-1995; 95US-0475886.

PF 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX WPI; 2000-586464/55.

DR N-PSDB; AAA75476.

PT Novel live hepatitis A virus adapted to growth in human fibroblast cell

PS line useful as vaccine for protecting humans against hepatitis A virus

XX infection, has modified genome compared to wild type

CC Disclosure; Fig 6A-K; 72pp; English.

CC The present sequence is derived from a wild type hepatitis A virus

CC (HAV) strain HM-174. The sequence is modified to produce HAV which

CC are adapted to growth in the human fibroblast-like cell line MRC-5.

CC The HAV is able to propagate in MRC-5 cells and retain appropriate

CC attenuation. It is useful as a live vaccine for prophylaxis of

CC hepatitis A in humans and other primates.

# RESULT 9

AAB18608  
 ID AAB18608 standard; Protein; 2227 AA.

AC AAB18608;

DT 15-JAN-2001 (first entry)

DE Amino acid sequence of passage 35 Hepatitis A virus called P-35.

KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;

XX P-35 virus.

OS Hepatitis A virus.

PN US6113912-A.

PD 05-SEP-2000.

PF 07-JUN-1995; 95US-0475886.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;

XX WPI; 2000-586464/55.

DR N-PSDB; AAA75477.

PT Novel live hepatitis A virus adapted to growth in human fibroblast cell

PS line useful as vaccine for protecting humans against hepatitis A virus

XX infection, has modified genome compared to wild type

CC Disclosure; Columns 67-78; 72pp; English.

CC The present sequence is derived from passage 35 of a wild type

CC hepatitis A virus (HAV) strain HM-174. The resulting virus is

CC designated P-35 virus. The sequence is modified to produce HAV which

CC are adapted to growth in the human fibroblast-like cell line MRC-5.

CC The HAV is able to propagate in MRC-5 cells and retain appropriate

CC attenuation. It is useful as a live vaccine for prophylaxis of

CC hepatitis A in humans and other primates.

CC

CC

CC

CC

CC

CC

CC

CC

CC

# RESULT 10

AAB18609 standard; Protein; 2227 AA.

ID AAB18609 standard; Protein; 2227 AA.

AC AAB18609;

DT 15-JAN-2001 (first entry)

DE Amino acid sequence of live attenuated Hepatitis A virus 4380.

KW HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;

XX HAV 4380.

OS Hepatitis A virus.

PN US6113912-A.

PD 05-SEP-2000.  
XX  
PF 07-JUN-1995; 95US-0475886.  
XX  
PR 18-SEP-1992; 92US-0947338.  
PR 17-SEP-1993; 93WO-US08610.  
PR 10-MAR-1995; 95US-0397232.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AM;  
XX  
DR WPI: 2000-586464/55.  
DR N-PSDB; AAA75478.  
XX  
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell  
PT line useful as vaccine for protecting humans against hepatitis A virus  
PT infection, has modified genome compared to wild type  
XX  
PS Disclosure; Columns 93-104; 72pp; English.  
XX  
XX The present sequence is derived from a live attenuated hepatitis A  
CC virus (HAV) of the invention, designated HAV 4380. The sequence is  
CC produced by modifying wild type HAV strain HM-174. The HAV of the  
CC invention are adapted to growth in the human fibroblast-like cell  
CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain  
CC appropriate attenuation. It is useful as a live vaccine for prophylaxis  
CC of hepatitis A in humans and other primates.  
XX  
SQ Sequence 2227 AA;  
XX  
XX Query Match 100.0%; Score 96; DB 21; Length 2227;  
XX Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
OY 1 KINLADRMIGLSGVOEIKQ 20  
DB 961 KINLADRMIGLSGVOEIKQ 980  
XX  
XX RESULT 11  
XX ABG31727  
XX ID ABG31727 standard; Protein; 2227 AA.  
XX AC ABG31727;  
XX DT 29-NOV-2002 (first entry)  
XX  
XX Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.  
XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.  
XX OS Hepatitis A virus strain HM-175.  
XX  
XX US6423318-B1.  
XX  
XX 23-JUL-2002.  
XX  
XX 31-AUG-2000; 2000US-0653499.  
XX  
XX 07-JUN-1995; 95US-0475886.  
XX PR 17-SEP-1993; 93US-0397232.  
XX PR 17-SEP-1993; 93WO-US08610.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.  
XX  
XX Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;  
XX  
XX WPI: 2002-680946/73.  
XX DR N-PSDB; ABB52787.  
XX  
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth

PT in MRC-5 cells, useful for preparing a vaccine against HAV infection  
XX  
XX Disclosure; Fig 6; 71pp; English.  
XX  
XX The invention relates to a polynucleotide which encodes a hepatitis A  
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell  
CC line). The polynucleotide is useful for preparing a vaccine against  
CC hepatitis A virus infection. This sequence represents a hepatitis A virus  
CC strain HM-175 polypeptide.  
XX  
SQ Sequence 2227 AA;  
XX  
XX Query Match 100.0%; Score 96; DB 23; Length 2227;  
XX Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
OY 1 KINLADRMIGLSGVOEIKQ 20  
DB 961 KINLADRMIGLSGVOEIKQ 980  
XX  
XX RESULT 12  
XX ABG31728  
XX ID ABG31728 standard; Protein; 2227 AA.  
XX AC ABG31728;  
XX DT 29-NOV-2002 (first entry)  
XX  
XX Hepatitis A virus mutant strain HM-175/7 (pHAV/7) polypeptide.  
XX  
XX Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;  
KW virucide; mutant; pHAV/7; mutcin.  
XX  
XX OS Hepatitis A virus strain HM-175.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
XX FH 963  
XX FT Misc-difference  
XX FT /Label= Wild-type Lys substituted by Arg  
XX FT Misc-difference 764  
XX FT /note= "Wild-type Glu substituted by Val"  
XX FT Misc-difference 821  
XX FT /note= "Wild-type Asn substituted by Ser"  
XX FT Misc-difference 1052  
XX FT /note= "Wild-type Ala substituted by Val"  
XX FT Misc-difference 1062  
XX FT /note= "Wild-type Gly substituted by Ala"  
XX FT Misc-difference 1118  
XX FT /note= "Wild-type Lys substituted by Met"  
XX FT Misc-difference 1151  
XX FT /note= "Wild-type Glu substituted by Lys"  
XX FT Misc-difference 1163  
XX FT /note= "Wild-type Phe substituted by Ser"  
XX FT Misc-difference 1277  
XX FT /note= "Wild-type Val substituted by Ile"  
XX FT Misc-difference 1500  
XX FT /note= "Wild-type His substituted by Tyr"  
XX FT Misc-difference 1805  
XX FT /note= "Wild-type Asp substituted by Asn"  
XX FT Misc-difference 1930  
XX FT /note= "Wild-type Ser substituted by Thr"  
XX  
XX US6423318-B1.  
XX  
XX 23-JUL-2002.  
XX  
XX 31-AUG-2000; 2000US-0653499.  
XX  
XX 07-JUN-1995; 95US-0475886.  
XX PR 17-SEP-1993; 93US-0397232.  
XX PR 17-SEP-1993; 93WO-US08610.  
XX

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA (SMK ) SMITHKLINE BECHAM BIOLOGICALS.  
PI Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;  
XX  
XX  
DR WPI: 2002-680946/73.  
DR N-PSDB; ABS52788.  
XX  
XX  
PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth  
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection -  
XX  
XX  
PS Example 3; Column 67-78; 71pp; English.  
XX  
XX  
CC The invention relates to a polynucleotide which encodes a hepatitis A  
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell  
CC line). The polynucleotide is useful for preparing a vaccine against  
CC hepatitis A virus infection. This sequence represents a hepatitis A virus  
CC mutant strain HM-175/77 (pHAV/77) polypeptide.  
XX  
XX  
Sequence 2227 AA:  
Query Match 100.0%; Score 96; DB 23; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 KINLADRMGLSGVGEIRKQ 20  
DB 961 KINLADRMGLSGVGEIRKQ 980  
RESULT 13  
ABG31729  
ID ABG31729 standard; Protein; 2227 AA.  
XX  
XX  
AC ABG31729;  
XX  
XX  
DT 29-NOV-2002 (first entry)  
XX  
XX  
DE Attenuated Hepatitis A virus (HAV) 4380 polypeptide.  
XX  
XX  
KM Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;  
KM HAV 4380.  
XX  
XX  
OS Hepatitis A virus strain HM-175.  
XX  
XX  
PN US6423318-B1.  
XX  
XX  
PS 23-JUL-2002.  
XX  
XX  
31-AUG-2000; 2000US-0653499.  
XX  
XX  
PR 07-JUN-1995; 95US-0475886.  
PR 17-SEP-1993; 93US-0397232.  
PR 17-SEP-1993; 93WO-US08610.  
XX  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PA (SMK ) SMITHKLINE BECHAM BIOLOGICALS.  
XX  
XX  
PI Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;  
XX  
XX  
DR WPI: 2002-680946/73.  
DR N-PSDB; ABS52789.  
XX  
XX  
PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth  
PT in MRC-5 cells, useful for preparing a vaccine against HAV infection -  
XX  
XX  
PS Disclosure; Column 93-104; 71pp; English.  
XX  
XX  
CC The invention relates to a polynucleotide which encodes a hepatitis A  
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell  
CC line). The polynucleotide is useful for preparing a vaccine against  
CC hepatitis A virus infection. This sequence represents an attenuated  
CC hepatitis A virus 4830 polypeptide.

XX  
SQ Sequence 2227 AA;  
Query Match 100.0%; Score 96; DB 23; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 KINLADRMGLSGVGEIRKQ 20  
DB 961 KINLADRMGLSGVGEIRKQ 980  
RESULT 14  
AAE19899  
ID AAE19899 standard; Protein; 2227 AA.  
XX  
XX  
AC AAE19899;  
XX  
XX  
DT 18-JUN-2002 (first entry)  
XX  
XX  
DE Hepatitis A virus (HAV) protein.  
XX  
XX  
KM Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;  
KM cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.  
XX  
XX  
OS Hepatitis A virus.  
XX  
XX  
PN WO200213855-A2.  
XX  
XX  
PD 21-FEB-2002.  
XX  
XX  
PF 15-AUG-2001; 2001WO-IB01808.  
XX  
XX  
PR 17-AUG-2000; 2000US-225767P.  
PR 29-AUG-2000; 2000US-229175P.  
PR 03-NOV-2000; 2000US-0705547.  
XX  
XX  
PA (TRIP-) TRIPEP AB.  
XX  
XX  
PI Salberg M, Hultgren C;  
XX  
XX  
DR WPI: 2002-241837/29.  
DR N-PSDB; AAD31766.  
XX  
XX  
PT Vaccine compositions for treating and preventing disease, preferably  
PT hepatitis C virus infection, comprises ribavirin and antigen that has  
PT epitope present in hepatitis C virus -  
XX  
XX  
PS Claim 11; Page 82-87; 120pp; English.  
XX  
XX  
CC The invention relates to a composition comprising ribavirin and an  
CC antigen preferably non structural 3 protein (NS3)/4A fragment of  
CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV  
CC sequence. The composition is useful for enhancing an immune response to  
CC a hepatitis C antigen in humans, domestic, sport or pet species and as  
CC vaccines for treating and preventing HCV infections. The composition is  
CC also useful for treating viral, bacterial, fungal diseases and cancer.  
CC The present sequence is hepatitis A virus (HAV) protein.  
XX  
XX  
SQ Sequence 2227 AA;  
Query Match 100.0%; Score 96; DB 23; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.2e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 KINLADRMGLSGVGEIRKQ 20  
DB 961 KINLADRMGLSGVGEIRKQ 980  
RESULT 15  
ABU08639  
ID ABU08639 standard; Protein; 2227 AA.

```

XX AB008639;
AC
XX
XX 03-JUN-2003 (first entry)
DT
XX
XX Wild type human hepatitis A virus strain HM-175.
DE
XX
XX Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;
KM vaccine; MRC-5 cell; hepatitis infection.
XX
XX Hepatitis A virus strain HM-175.
OS
XX
XX US2002176869-A1.
XX
XX
XX 28-NOV-2002.
XX
XX
XX 29-APR-2002; 2002US-0135988.
XX
XX
XX 07-JUN-1995; 95US-0475886.
XX 31-AUG-2000; 2000US-0653499.
XX 18-SEP-1992; 92US-0947338.
XX 17-SEP-1993; 93WO-US08610.
XX 17-APR-1995; 95US-0397232.
XX
XX
XX (FUNK/) FUNKHOUSER A W.
XX (EMER/) EMERSON S U.
XX (PURC/) PURCELL R H.
XX (DHON/) D'HONDT E.
XX
XX
XX Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;
XX
XX WPI: 2003-352605/33.
XX DR N-PSDB; ABX93473.
XX
XX New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,
XX useful in vaccines for protecting primates against hepatitis infection
XX and disease
XX
XX Disclosure: Fig 6; 70pp; English.
XX
XX
XX The invention describes a live hepatitis A virus (HAV) adapted to growth
XX in MRC-5 cells. The HAV and compositions comprising the HAV are useful as
XX a vaccine for protecting primates against hepatitis infection and
XX disease. This is the amino acid sequence of wild type human hepatitis A
XX virus strain HM-175.
XX
XX
XX Sequence 2227 AA;
XX
XX
XX Query Match 100.0%; Score 96; DB 24; Length 2227;
XX Best Local Similarity 100.0%; Pred. NO. 2.2e-07;
XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX 1 KINLADRMIGLSGVOEIREQ 20
XX ||||||||||||||||
XX 961 KINLADRMIGLSGVOEIREQ 980

```

Search completed: October 1, 2003, 09:56:45  
 Job time : 44.5294 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 12.8235 seconds  
(without alignments)  
149.988 Million cell updates/sec

Title: US-09-171-432A-46

Perfect score: 96

Sequence: 1 KINLADRMGLSGVOEIKRQ 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	2227	1 GNNYHM	genome polypotein
2	96	100.0	2227	1 GNNYHR	genome polypotein
3	96	100.0	2227	1 GNNYMK	genome polypotein
4	96	100.0	2227	1 GNNYHB	genome polypotein
5	96	100.0	2230	1 GNNYSA	genome polypotein
6	47	49.0	145	2 B25199	heat shock 16k pro
7	47	49.0	172	2 S33416	heat shock protein
8	44	45.8	414	2 T06303	enoyl-CoA hydratase
9	44	45.8	646	2 S72609	GTP-binding membra
10	44	45.8	653	2 G70683	probable lepa - My
11	44	45.8	788	2 S67595	hypothetical prote
12	43	44.8	145	2 B24289	heat shock protein
13	43	44.8	158	2 S64321	hypothetical prote
14	43	44.8	180	2 AC0056	hypothetical prote
15	43	44.8	370	2 H70423	conserved hypotet
16	43	44.8	2104	2 D91286	oxygen-independent
17	43	44.8	2104	2 H86127	hypothetical prote
18	42.5	44.3	602	2 AB3542	hypothetical prote
19	42	43.8	115	2 B86882	glp-binding protei
20	42	43.8	143	1 HHRW41	hypothetical prote
21	42	43.8	547	2 A40656	heat shock protein
22	42	43.8	547	2 AE1022	probable membrane
23	42	43.8	646	2 D81674	conserved hypotet
24	42	43.8	1035	2 S18512	cell division cont
25	42	43.8	1250	2 A39578	SSD1 protein - yea
26	41	42.7	141	2 E64368	hypothetical prote
27	41	42.7	155	2 C34965	hypothetical 17k p
28	41	42.7	155	2 S15576	lipid protein - Shi
29	41	42.7	231	2 AE2986	conserved hypotet

30	41	42.7	231	2 B98297	hypothetical prote
31	41	42.7	313	2 D95866	probable transcrip
32	41	42.7	366	1 TDMSM4	monocyte surface g
33	41	42.7	518	1 S44183	thiamine-phosphate
34	41	42.7	605	2 D83007	regulatory protein
35	41	42.7	621	2 A71516	hypothetical prote
36	41	42.7	687	2 T09051	Pepp protein - Pae
37	41	42.7	708	2 JG6329	yeast secretory pr
38	41	42.7	1036	2 F83263	probable multidrug
39	41	42.7	1238	2 T40993	protein kinase cck
40	41	42.7	1652	2 I50711	complement C3 prec
41	41	41.7	143	1 HHRW48	heat shock protein
42	40	41.7	208	2 A70122	glucose inhibited
43	40	41.7	247	2 F90656	hypothetical prote
44	40	41.7	247	2 F85507	hypothetical prote
45	40	41.7	332	2 JG4535	urate oxidase (EC

## ALIGNMENTS

## RESULT 1

GNNYHM  
N: genome polypotein - human hepatitis A virus (strain HM-175, wild type)  
B: N: contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core  
C: Species: human hepatitis A virus  
A: Note: host Homo sapiens (man)  
C: Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 16-Jul-1999  
C: Accession: A25981  
R: Cohen, J.L.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.  
A: Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with  
A: Reference number: A25981; MUID:87061253; PMID:3023706  
A: Accession: A25981  
A: Molecule type: genomic RNA  
A: Residues: 1-2227 <COH>  
A: Cross-references: EMBL:M14707; NID:q329582; PIDN:AAA45465.1; PID:q329583  
C: Superfamily: hepatitis A virus genome polypotein  
C: Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr  
F: 1-23/Product: coat protein 1A #status predicted <VP4>  
F: 24-245/Product: coat protein 1B #status predicted <VP2>  
F: 246-491/Product: coat protein 1C #status predicted <VP3>  
F: 492-791/Product: coat protein 1D #status predicted <VP1>  
F: 792-980/Product: core protein 2A #status predicted <C2A>  
F: 981-1087/Product: core protein 2B #status predicted <C2B>  
F: 1088-1422/Product: core protein 2C #status predicted <C2C>  
F: 1423-1496/Product: protein 3A #status predicted <C3A>  
F: 1497-1519/Product: protein 3B #status predicted <C3B>  
F: 1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>  
F: 1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 96; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2.8e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 961 KINLADRMGLSGVOEIKRQ 980

RESULT 2  
GNNYHR  
N: genome polypotein - human hepatitis A virus  
N: contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core  
NA polymerase (EC 2.7.7.48), protein 3D  
C: Species: human hepatitis A virus  
A: Note: host Homo sapiens (man)  
C: Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Jul-1999  
C: Accession: A03903  
R: Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N  
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985  
A: Title: Primary structure and gene organization of human hepatitis A virus.



F:28-249/Product: coat protein 1B #status predicted <C1B>  
F:250-495/Product: coat protein 1C #status predicted <C1C>  
F:496-795/Product: coat protein 1D #status predicted <C1D>  
F:796-984/Product: core protein 2A #status predicted <C2A>  
F:985-1091/Product: core protein 2B #status predicted <C2B>  
F:1092-1426/Product: core protein 2C #status predicted <C2C>  
F:1427-1498/Product: protein 3A #status predicted <P3A>  
F:1499-1521/Product: protein 3B #status predicted <P3B>  
F:1522-1741/Product: protein 3C #status predicted <P3C>  
F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 100.0%; Score 96; DB 1; Length 2230;  
Best Local Similarity 100.0%; Pred. No. 2.8e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKRQ 20  
Db 965 KINLADRMGLSGVOEIKRQ 984

## RESULT 6

heat shock 16K protein 2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 16-Aug-1988 #sequence\_revision 16-Aug-1988 #text\_change 13-Aug-1999  
C:Accession: B25199  
R:Jones, D.; Russnak, R.H.; Kay, R.J.; Candido, E.P.M.  
J. Biol. Chem. 261, 12006-12015, 1986  
A:Title: Structure, expression, and evolution of a heat shock gene locus in Caenorhabditis  
A:Reference number: A92555; MID:86304344; PMID:3017958  
A:Accession type: B25199  
A:Molecule type: DNA  
A:Residues: 1-145 <CON>  
A:Cross-references: GB:M14334; NID:g156338; PIDN:AAA28071.1; PID:g156340  
C:Superfamily: alpha-crystallin

Query Match 49.0%; Score 47; DB 2; Length 145;  
Best Local Similarity 55.6%; Pred. No. 2.2;  
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKR 18  
Db 65 KINLADRMGLSGVOEIKR 82

## RESULT 7

heat shock protein hsp20 - nematode (Nippostrongylus brasiliensis)  
C:Species: Nippostrongylus brasiliensis  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 13-Aug-1999  
C:Accession: S33416  
R:Tweedie, S.; Grieg, M.E.; Ingram, L.; Selkirk, M.E.  
submitted to the EMBL Data Library, April 1993  
A:Description: The expression of a small heat shock homologue is developmentally regulated  
A:Reference number: S33416  
A:Accession: S33416  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-172 <TWB>  
A:Cross-references: EMBL:X71663; NID:g297865; PIDN:CAA50655.1; PID:g297866  
C:Superfamily: alpha-crystallin

Query Match 49.0%; Score 47; DB 2; Length 172;  
Best Local Similarity 45.0%; Pred. No. 2.6;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKRQ 20  
Db 80 KVALDRDLTVBGMDEVKTE 99

## RESULT 8

T06303

enoyl-CoA hydratase homolog F11C18.10 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 22-Oct-1999  
C:Accession: T06303  
R:Bevan, M.; Terry, N.; Ardiles, W.; Blyshaett, C.; Dasseville, R.; De Clerck, R.;  
ewes, H.W.; Mayer, K.F.X.; Schueller, C.  
submitted to the Protein Sequence Database, April 1999  
A:Reference number: Z15589  
A:Accession: T06303  
A:Molecule type: DNA  
A:Residues: 1-414 <REV>

A:Cross-references: EMBL:AL049607; GSPDB:GN00062; ATSP:F11C18.10  
A:Experimental source: cultivar Columbia; BAC clone F11C18  
C:Genetics:  
A:Gene: ATSP:F11C18.10  
A:Map position: 4  
A:Introns: 42/3; 75/2; 108/1; 134/3; 162/3; 191/1; 219/3; 260/2; 280/3; 313/3; 347/3;

Query Match 45.8%; Score 44; DB 2; Length 414;  
Best Local Similarity 52.6%; Pred. No. 22;  
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKR 19  
Db 260 KIELDKYRGLDVEIKR 278

## RESULT 9

GTP-binding membrane protein lepa - Mycobacterium leprae  
S72609  
N:Alternate names: protein B1937\_f3\_81  
C:Species: Mycobacterium leprae  
C:Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 02-Feb-2001  
C:Accession: S72609  
R:Smith, D.R.; Robison, K.  
submitted to the EMBL Data Library, November 1993  
A:Reference number: S72580  
A:Accession: S72609  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-646 <SMI>  
A:Cross-references: EMBL:U00016; NID:g466961; PIDN:AAA17177.1; PID:g466991  
C:Genetics:  
A:Gene: lepa  
A:Start codon: GNG  
A:Superfamily: GTP-binding membrane protein lepa; translation elongation factor Tu ho  
C:Keywords: GTP binding; membrane protein; nucleotide binding; P-loop  
F:45-177/Domain: translation elongation factor Tu homology <ETU>  
F:51-58/Region: nucleotide-binding motif A (P-loop)  
F:174-177/Region: GTP-binding NKXD motif

Query Match 45.8%; Score 44; DB 2; Length 646;  
Best Local Similarity 71.4%; Pred. No. 35;  
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGV 14  
Db 57 KSLADRMGLSGV 70

## RESULT 10

probable lepa - Mycobacterium tuberculosis (strain H37RV)  
G70683  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 02-Feb-2001  
C:Accession: G70683  
R:Coyle, S.T.; Brosch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
Rajandream, M.A.; Davies, R.; Devlin, K.; Felzwell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
A:Reference number: A70500; MID:98295987; PMID:9634230

A:Accession: G70683  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-653 <COL>  
A:Cross-references: GB:281368; GB:AL123456; NID:g3261656; PIDN:CAB03723.1; PID:g1655655  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: LepA  
C:Superfamily: GTP-binding membrane protein LepA; translation elongation factor Tu homology  
C:Keywords: GTP binding; nucleotide binding; P-loop  
F:53-181/Domain: translation elongation factor Tu homology <ETU>  
F:59-66/Region: nucleotide-binding motif A (P-loop)  
F:178-181/Region: GTP-binding NKXD motif

Query Match 45.8%; Score 44; DB 2; Length 653;  
Best Local Similarity 71.4%; Pred. No. 36;  
Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KINADRMGLSGV 14  
| | | | | | | | | | | | | | | | | |  
65 KSTLADRMLOLTGV 78

RESULT 11  
S67595  
hypothetical protein YDL060w - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein D2544  
C:Species: Saccharomyces cerevisiae  
C:Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 19-Apr-2002  
C:Accession: S67595  
R:Biocheck: H.; Brandt, P.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67587  
A:Accession: S67595  
A:Molecule type: DNA  
A:Residues: 1-788 <BIO>  
A:Cross-references: EMBL:Z74108; NID:g1431062; PID:g1431063; GSPDB:GN00004; MIPS:YDL060w  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:TSR1; MIPS:YDL060w  
A:Cross-references: SGD:S0002218  
A:Map position: 4L

Query Match 45.8%; Score 44; DB 2; Length 788;  
Best Local Similarity 61.5%; Pred. No. 44;  
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 8 MLGSGVOEIREQ 20  
: | | | | | | | | | | | | | | | | | |  
159 VFGLSGVOEYDEE 171

RESULT 12  
B24289  
heat shock protein 16-1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 28-Sep-1987 #sequence\_revision 30-Sep-1993 #text\_change 29-Oct-1999  
C:Accession: B24289; S31037; T25927; T25930  
R:Rusnak, R.H.; Candido, E.P.M.  
Mol. Cell. Biol. 5, 1268-1278, 1985  
A:Reference number: A24289; MUID:85295957; PMID:4033652  
A:Accession: B24289  
A:Molecule type: DNA  
A:Residues: 1-145 <RUS>  
A:Cross-references: EMBL:K03273; NID:g156333; PIDN:AAA28068.1; PID:g156336  
A>Note: the author translated the codon GAT for residue 17 as Tyr  
R:Kay, R.U.; Rusnak, R.H.; Jones, D.; Mathias, C.; Candido, E.P.M.  
Nucleic Acids Res. 15, 3723-3741, 1987  
A:Title: Expression of intron-containing C. elegans heat shock genes in mouse cells dem-  
o Y effect of heat shock on the mammalian splicing apparatus.  
A:Reference number: S31036; MUID:87231065; PMID:3588308  
A:Accession: S31037  
A:Status: translation not shown

A:Molecule type: DNA  
A:Residues: 41-63 <KAY>  
A:Cross-references: EMBL:M31340; NID:g156343; PIDN:AAA28073.1; PID:g552068  
R:Bradshaw, H.  
submitted to the EMBL Data Library, July 1996  
A:Description: The sequence of C. elegans cosmid T27E4.  
A:Reference number: Z20111  
A:Accession: T25927  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-145 <BR2>  
A:Cross-references: EMBL:U64837; PIDN:AB04839.1; GSPDB:GN00023; CESP:hsp-16A  
A:Experimental source: strain Bristol N2; clone T27E4  
A:Accession: T25930  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-145 <BR2>  
A:Cross-references: EMBL:U64837; PIDN:AB04842.1; GSPDB:GN00023; CESP:hsp-16A  
A:Experimental source: strain Bristol N2; clone T27E4  
C:Genetics:  
A:Gene: CESP:hsp-16A  
A:Map position: 5  
A:Introns: 42/3  
C:Superfamily: alpha-crystallin

Query Match 44.8%; Score 43; DB 2; Length 145;  
Best Local Similarity 45.0%; Pred. No. 10;  
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 1 KINADRMGLSGVOEIREQ 20  
| | | | | | | | | | | | | | | | | |  
65 KINLDGHTLSIGEOELKTE 84

RESULT 13  
S64321  
hypothetical protein YGR030c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein G4068  
C:Species: Saccharomyces cerevisiae  
C:Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 19-Apr-2002  
C:Accession: S64321  
R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.  
submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64071  
A:Accession: S64321  
A:Molecule type: DNA  
A:Residues: 1-158 <RIE>  
A:Cross-references: EMBL:Z72815; NID:g1323007; PID:e243932; PID:g1323008; GSPDB:GN000  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:POP6; MIPS:YGR030c  
A:Cross-references: SGD:S0003262  
A:Map position: 7R

Query Match 44.8%; Score 43; DB 2; Length 158;  
Best Local Similarity 64.3%; Pred. No. 11;  
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 KINADRMGLSGV 14  
: | | | | | | | | | | | | | | | | | |  
65 QINMADRSLGLQOV 78

RESULT 14  
AC0056  
conserved hypothetical protein YP00454 [Imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001  
C:Accession: AC0056  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.  
demo-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.  
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell  
Nature 413, 523-527, 2001



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen, Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 ; Search time 6.58824 Seconds

(Without alignments)  
142.760 Million cell updates/sec

Title: US-09-171-432a-46  
Perfect score: 96

Sequence: 1 KINADRMGLSGVQEKREQ 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Minimum number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	2226	1	POLG_HPAV2
2	96	100.0	2226	1	POLG_HPAV4
3	96	100.0	2226	1	POLG_HPAV8
4	96	100.0	2227	1	POLG_HPAVH
5	96	100.0	2227	1	POLG_HPAVH
6	96	100.0	2227	1	POLG_HPAVH
7	96	100.0	2230	1	POLG_HPAVH
8	48	50.0	622	1	LEPA_STRCO
9	47	49.0	145	1	HS12_CAEEL
10	47	49.0	172	1	HS12_NIPBR
11	44	45.8	646	1	LEPA_MYCLE
12	44	45.8	653	1	LEPA_MYCTU
13	43	44.8	145	1	HS11_CAEEL
14	43	44.8	158	1	POPE_YEAST
15	43	44.8	180	1	Y454_YERPE
16	42.5	44.3	602	1	LEPA_BDOME
17	42	43.8	143	1	HS16_CAEEL
18	42	43.8	151	1	SODC_HALRO
19	42	43.8	547	1	YDDB_SALTY
20	42	43.8	621	1	Y708_CHLMY
21	42	43.8	1035	1	CC68_YEAST
22	42	43.8	1250	1	SSD1_YEAST
23	41	42.7	141	1	NIRK_METUA
24	41	42.7	155	1	IPPI_SHIFL
25	41	42.7	313	1	CBR1_RHIME
26	41	42.7	313	1	CBR2_RHIME
27	41	42.7	366	1	CD14_MOUSE
28	41	42.7	518	1	ASB6_MOUSE
29	41	42.7	518	1	TH14_SCHPO
30	41	42.7	621	1	Y425_CHLTP
31	41	42.7	708	1	SC10_HUMAN
32	41	42.7	708	1	SC10_RAT
33	41	42.7	1338	1	CEK1_SCHPO

34	40	41.7	143	1	HS17_CAEEL	P02513
35	40	41.7	208	1	GIDB_BORBU	P53363
36	40	41.7	316	1	OCDB_HUMAN	P09417
37	40	41.7	502	1	URIC_BACSB	Q45697
38	40	41.7	615	1	LEPA_BACSL	Q8nn68
39	40	41.7	698	1	MCHE_ECOLI	Q8nn68
40	40	41.7	739	1	RELA_STREQ	Q54089
41	40	41.7	819	1	LOH_CHLPN	Q54089
42	40	41.7	1042	1	SVI_BORBU	Q54174
43	40	41.7	1227	1	CIBE_BACTV	Q51773
44	40	41.7	1228	1	CIBA_BACTV	Q51773
45	40	41.7	1229	1	CIBB_BACTV	Q45739

## ALIGNMENTS

RESULT 1  
POLG\_HPAV2 STANDARD: PRT: 2226 AA.  
ID POLG\_HPAV2  
AC P26580;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain 24a).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
ON NCBI\_TaxID=12094;  
RX MEDLINE=91162758; PubMed=1705995;  
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,  
RA "Antigenic and genetic variation in cytopathic hepatitis A virus  
RT variants arising during persistent infection: evidence for genetic  
RT recombination.";  
RL J. Virol. 65:2056-2065(1991).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
(RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sb.ch/announce/>  
or send an email to [license@isb-sb.ch](mailto:license@isb-sb.ch)).  
CC EMBL: M59810; AAA5468.1; -.  
DR MEROPS: C03.005; -.  
DR InterPro: IPR004004; Calic1\_pol\_hel.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR007094; RNA\_dep\_RNA\_pol.  
DR pfam: PF00680; RNA\_dep\_RNA\_pol. 1.  
DR pfam: PF00910; RNA\_helicase. 1.  
DR PRINTS: PR00918; CALICVIRUSNS.  
KW polyprotein; Coat protein; Core protein; Core protein; Transferase;  
RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
KW CHAIN 1  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 794  
FT CHAIN 795 900  
CORE PROTEIN P2A.

FT CHAIN 901 1087 CORE PROTEIN P2B.  
FT CHAIN 1088 1422 CORE PROTEIN P2C.  
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.  
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.  
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.  
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.  
SQ SEQUENCE 2226 AA; 251152 MM; 6CD85A91D6B4E2BF CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2226;  
Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20  
DB 961 KINLADRMGLSGVQEIKEQ 980

RESULT 2  
POLG\_HPAV4 STANDARD; PRT; 2226 AA.

P26581;  
01-AUG-1992 (Rel. 23, Created)

01-AUG-1992 (Rel. 23, Last sequence update)

28-FEB-2003 (Rel. 41, Last annotation update)

Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (PC 2.7.7.48)].

Hepatitis A virus (strain 43c).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.

NCBI\_Taxid=12095;

SEQUENCE FROM N.A.  
MEDLINE=9162756; PubMed=1705995;

Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M., Cromeans T., Jansen R.W.;

"Antigenic and genetic variation in cytopathic hepatitis A virus RT variants arising during persistent infection: evidence for genetic recombination."

RT J. Virol. 65:2056-2065(1991).

-1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + [RNA](N).

-1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.

-1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.1sb-sib.ch/announce/> or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).

EMBL; M59809; AAA45469.1; -

DR MEROPS; C03.005; -

DR InterPro; IPR000605; RNA\_helicase.

DR InterPro; IPR007095; RNA\_pol\_DS\_PS.

DR InterPro; IPR001205; RNA\_pol\_P3D.

DR InterPro; IPR007094; RNA\_pol\_PSVir.

DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.

DR Pfam; PF00910; RNA\_helicase; 1.

KW Polyprotein; Coat protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolyase; Thiol protease.

FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).

FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).

FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).

FT CHAIN 492 794 COAT PROTEIN VP1 (PID).

FT CHAIN 795 900 CORE PROTEIN P2A.

FT CHAIN 901 1087 CORE PROTEIN P2B.

FT CHAIN 1088 1422 CORE PROTEIN P2C.

FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.  
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.  
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.  
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.  
SQ SEQUENCE 2226 AA; 251107 MM; 403B4CA80B95F75 CRC64;

Query Match 100.0%; Score 96; DB 1; Length 2226;  
Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVQEIKEQ 20  
DB 961 KINLADRMGLSGVQEIKEQ 980

RESULT 3  
POLG\_HPAV8 STANDARD; PRT; 2226 AA.

P26582;  
01-AUG-1992 (Rel. 23, Created)

01-AUG-1992 (Rel. 23, Last sequence update)

15-SEP-2003 (Rel. 42, Last annotation update)

Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (PC 2.7.7.48)].

Hepatitis A virus (strain 18f).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Hepatovirus.

NCBI\_Taxid=12096;

SEQUENCE FROM N.A.  
MEDLINE=9162756; PubMed=1705995;

Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M., Cromeans T., Jansen R.W.;

"Antigenic and genetic variation in cytopathic hepatitis A virus RT variants arising during persistent infection: evidence for genetic recombination."

RT J. Virol. 65:2056-2065(1991).

-1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + [RNA](N).

-1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.

-1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.

-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.1sb-sib.ch/announce/> or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).

EMBL; M59808; AAA45467.1; -

DR PDB; 1QAV; 15-MAY-00.

DR MEROPS; C03.005; -

DR InterPro; IPR000605; RNA\_helicase.

DR InterPro; IPR007095; RNA\_pol\_DS\_PS.

DR InterPro; IPR001205; RNA\_pol\_P3D.

DR InterPro; IPR007094; RNA\_pol\_PSVir.

DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.

DR Pfam; PF00910; RNA\_helicase; 1.

KW Polyprotein; Coat protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolyase; Thiol protease; 3D-structure.

FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).

FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).

FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).

FT CHAIN 492 794 COAT PROTEIN VP1 (PID).

FT CHAIN 795 900 CORE PROTEIN P2A.

FT CHAIN 901 1087 CORE PROTEIN P2B.

FT CHAIN 1088 1422 CORE PROTEIN P2C.

FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.





RT "Primary structure and gene organization of human hepatitis A virus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC {RNA}(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: K02990; AAA45472.1; -.  
CC PIR: A03903; GNNTHR.  
CC -----  
CC InterPro: IPR004004; Calic1\_pol\_hel.  
CC InterPro: IPR000605; RNA\_helicase.  
CC InterPro: IPR007095; RNA\_pol\_DS\_PS.  
CC InterPro: IPR001205; RNA\_pol\_P3D.  
CC InterPro: IPR007094; RNA\_pol\_PSVlr.  
CC Pfam: PF00680; RNA\_dep\_RNA\_pol\_1.  
CC Pfam: PF00910; RNA\_helicase; 1.  
CC PRINTS: PR00918; CALICIVIRUSNS.  
CC -----  
CC Polyprotein: Coat protein; Core protein; Transferase;  
CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease.  
CC -----  
CC CHAIN 1 23  
CC FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).  
CC FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).  
CC FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).  
CC FT CHAIN 492 836 COAT PROTEIN VP1 (PID).  
CC FT CHAIN 837 980 CORE PROTEIN P2A.  
CC FT CHAIN 981 1076 CORE PROTEIN P2B.  
CC FT CHAIN 1077 1422 CORE PROTEIN P2C.  
CC FT CHAIN 1423 1484 PROBABLE PROTEIN P3A.  
CC FT CHAIN 1485 1507 PROBABLE PROTEIN P3B.  
CC FT CHAIN 1508 1678 PROBABLE PROTEIN P3C.  
CC FT CHAIN 1679 2227 RNA-DIRECTED POLYMERASE P3D.  
CC FT CHAIN 2227 AA; 251898 MW; 99A7354B4CD2799C CR64;  
CC -----  
CC Query Match 100.0%; Score 96; DB 1; Length 2227;  
CC Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
CC Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CC -----  
CC 1 KINLADRLGLSGVOETKEQ 20  
CC |||||||||||||||||||  
CC 961 KINLADRLGLSGVOETKEQ 980

RT "The entire nucleotide sequence of the genome of human hepatitis A  
RT virus (isolate MB8).";  
RL Virus Res. 8:153-171(1987).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC {RNA}(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT  
CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERANEAN REGION.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: M20273; AAA45474.1; -.  
CC PIR: A03903; GNNTHR.  
CC -----  
CC InterPro: IPR000605; RNA\_helicase.  
CC InterPro: IPR007095; RNA\_pol\_DS\_PS.  
CC InterPro: IPR001205; RNA\_pol\_P3D.  
CC InterPro: IPR007094; RNA\_pol\_PSVlr.  
CC Pfam: PF00680; RNA\_dep\_RNA\_pol\_1.  
CC Pfam: PF00910; RNA\_helicase; 1.  
CC PRINTS: PR00918; CALICIVIRUSNS.  
CC -----  
CC Polyprotein: Coat protein; Core protein; Transferase;  
CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease.  
CC -----  
CC CHAIN 1 23  
CC FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).  
CC FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).  
CC FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).  
CC FT CHAIN 492 836 COAT PROTEIN VP1 (PID).  
CC FT CHAIN 837 980 CORE PROTEIN P2A.  
CC FT CHAIN 981 1087 CORE PROTEIN P2B.  
CC FT CHAIN 1088 1422 CORE PROTEIN P2C.  
CC FT CHAIN 1423 1496 PROBABLE PROTEIN P3A.  
CC FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.  
CC FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.  
CC FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.  
CC FT CHAIN 2227 AA; 251425 MW; EC983E2A7C86449 CR64;  
CC -----  
CC Query Match 100.0%; Score 96; DB 1; Length 2227;  
CC Best Local Similarity 100.0%; Pred. No. 2.1e-07;  
CC Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CC -----  
CC 1 KINLADRLGLSGVOETKEQ 20  
CC |||||||||||||||||||  
CC 961 KINLADRLGLSGVOETKEQ 980

RESULT 7  
POLG\_HPAYS STANDARD: PRT: 2230 AA.  
ID POLG\_HPAYS  
AC P14553;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins  
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
DE P3D (EC 2.7.7.48)].  
DE P3D (EC 2.7.7.48)].  
OS Simian hepatitis A virus (strain AGM-27).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
CC NCBI\_TaxID=12102;  
CC [1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=91311420; PubMed=1649901;  
RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,  
RA Purcell R.H.;  
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome

**us-09-171-432a-46.rsp**

CC Streptomyces; Streptomycetaceae; Streptomycetes.  
OX NCBI\_TaxID=1902;  
  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2) / M145;  
RX MEDLINE=21996410; PubMed=12000953;  
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
RA Thompson N.R., James K.D., Harris D.E., Quall M.A., Kieser H.,  
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
RA Cronin A., Fraser A., Gobie A., Hidalgo J., Hornsby T., Howarth S.,  
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
RA Rhabdlowitsch E., Rajandream M.A., Rutherford K., Ruter S.,  
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
RA Hopwood D.A.;  
RT "Complete genome sequence of the model actinomycete Streptomyces  
RT coelicolor A3(2).";  
RL Nature 417:141-147(2002).  
CC -1 SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY,  
CC LEPA SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb.ch/announce/  
CC or send an email to license@isb.slb.ch).  
-----  
DR EMBL, AL939113, CAB66240.1; -.  
DR HSSP, P13551; ZEEG.  
DR HAMAP, MF\_00071; ? 1.  
DR InterPro: IPR000795; EF\_Grpbind.  
DR InterPro: IPR000640; EEG\_C.  
DR InterPro: IPR004161; EFFU\_D2.  
DR InterPro: IPR006297; LepA.  
DR InterPro: IPR005225; Small\_GTP.  
DR Pfam, PF00679; EEG\_C; 1.  
DR Pfam, PF00009; GTP\_EFTU; 1.  
DR Pfam, PF03144; GTP\_EFTU\_D2; 1.  
DR PRINTS: PRO0315; ELONGATNFCT.  
DR TRGFAMS: TGFR01393; LepA; 1.  
DR TRGFAMS: TGCR00231; small\_GTP; 1.  
DR PROSITE, PS00301; EFACTOR\_GTP; 1.  
KW GTP-binding; Complete proteome.  
FT NP\_BIND 26 GTP (BY SIMILARITY).  
FT NP\_BIND 94 GTP (BY SIMILARITY).  
FT NP\_BIND 148 151 GTP (BY SIMILARITY).  
FT FT 622 AA; 68378 MW; 83f5c76fr2ab0ccC CNC64;  
SQ SEQUENCE

OY 1 KINADRMALIGSGVOEIKE 19  
Db | ||||| :| :: 50  
32 KSTLDARMQLTGVEQRO 50

Query Match 50.0%; Score 48; DB 1; Length 622;  
Best Local Similarity 57.9%; Pred. No. 4.1;  
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

RESULT 9  
HS12\_CAEBL  
AC ID HS12\_CAEBL STANDARD; PRN; 145 AA.  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 28-FEB-2003 (Rel. 01, last annotation update)  
DE Heat shock protein HSP16-2.  
GN HSP16-2 OR Y46HA3.3.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;  
OC Rhabditiidae; Peloiderinae; Caenornaditils.  
OX NCBI\_TaxID=6239;  
NN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE-6304344; PubMed-3017958;  
RA Jones D., Russnak R.H., Kay R.J., Candido E.P.M.;  
RT "Structure, expression, and evolution of a heat shock gene locus in  
RT *Caenorhabditis elegans* that is flanked by repetitive elements.";   
RL J. Biol. Chem. 261:12006-12015(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Bristol N2;  
RA Stoneking T., Wohlmann P., Lennox S.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP REVISIONS.  
RA Waterston R.;  
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)  
FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC  
CC EMBL: M14334; AAA28071.1; -;  
DR EMBL: AC006774; AAF60615.1; -;  
DR PIR: B25199; B25199.  
DR WormPep: Y46H3A.3; CE22002.  
DR InterPro: IPR002068; Hsp20.  
DR Pfam: PF00011; HSP20; 1.  
DR PROSITE: PS01031; HSP20; 1.  
DR Heat shock; Multigene family.  
KW SEQUENCE 145 AA; 16242 MW; 8A73449F99161889 CRC64;  
SQ  
Query Match 49.0%; Score 47; DB 1; Length 145;  
Best Local Similarity 55.6%; Pred. No. 1.3;  
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
OY 1 KINLADRLGSGVQEI 18  
ID 65 KINLADRLGSGVQEI 82  
Db  
RESULT 10  
HSP20\_NIPBR STANDARD; PRT; 172 AA.  
ID 007160;  
01-OCT-1994 (Rel. 30, Created)  
01-OCT-1994 (Rel. 30, Last sequence update)  
01-OCT-1994 (Rel. 30, Last annotation update)  
DE Heat shock protein homolog (HSP20).  
GN HSP20.  
OS *Nippostrongylus brasiliensis*.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;  
OC Trichostrongylidae; Heliomonellidae; Nippostrongylinae;  
OC Nippostrongylus.  
OX NCBI\_TaxID=27835;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-94081860; PubMed-8259127;  
RA Tweedle S., Grlsog M.E., Ingram L., Selkirk M.E.;  
RT "The expression of a small heat shock protein homologue is  
RT developmentally regulated in *Nippostrongylus brasiliensis*.";   
RL Mol. Biochem. Parasitol. 61:149-154(1993).  
CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)  
FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC  
CC EMBL: X71663; CA50655.1; -;  
DR PIR: S33416; S33416.  
DR InterPro: IPR001436; Crystallin\_alpha.  
DR InterPro: IPR002068; Hsp20.  
DR Pfam: PF00011; HSP20; 1.  
DR PRINTS: PR00299; ACRYSTALLIN.  
DR PROSITE: PS01031; HSP20; 1.  
DR Heat shock.  
KW SEQUENCE 172 AA; 20227 MW; 2CDA711CE60B1C0 CRC64;  
SQ  
Query Match 49.0%; Score 47; DB 1; Length 172;  
Best Local Similarity 45.0%; Pred. No. 1.6;  
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
OY 1 KINLADRLGSGVQEI 20  
ID 80 KVLQDDRLDVEGMEVKE 99  
Db  
RESULT 11  
LEPA\_MTCLE STANDARD; PRT; 646 AA.  
ID LEPA\_MTCLE  
AC P53530;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE GTP-binding protein lepa.  
GN LEPA OR ML0611 OR B1937\_F3\_81.  
OS *Mycobacterium leprae*.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1769;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Smith D.R., Robison K.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TN;  
RX MEDLINE-21128732; PubMed-11234002;  
RA Cole S.T., Elgineter K., Parkhill J., James K.D., Thomson N.R.,  
RA Wheeler P.R., Honore N., Garner T., Churcher C., Harris D.,  
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R.M., Devlin K., Duthoy S., Felkell T., Fraser A., Hamlin N.,  
RA Holroyd S., Hornsby T., Jagers K., Lacroix C., Maclean J., Moule S.,  
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
RA Barrell B.G.;  
RT "Massive gene decay in the leprosy bacillus."  
RL Nature 409:1007-1011(2001).  
CC -1- MISCELLANEOUS: THE OVERPRODUCTION OF THIS PROTEIN IS LETHAL TO  
CC M. LEPAE.  
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
LEPA SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC  
CC EMBL: U00016; AAA17177.1; -;  
DR EMBL: AL583919; CAC30119.1; -;  
DR PIR: S72609; S72609.  
DR HSP: P13551; 2EFG.  
DR Leproma; ML0611; -;

DR HAMAP: MF\_00071: -- 1.  
 DR InterPro: IPR000795; EF\_GTPbind.  
 DR InterPro: IPR000640; EFG\_C.  
 DR InterPro: IPR004161; EFTU\_D2.  
 DR InterPro: IPR006297; LepA.  
 DR InterPro: IPR005225; Small\_GTP.  
 DR Pfam: PF00679; EFG\_C.1.  
 DR Pfam: PF00009; GTP\_EFTU.1.  
 DR Pfam: PF03144; GTP\_EFTU\_D2.1.  
 DR PRINTS: PR00315; ELONGATNFCF.  
 DR TIGRPFAM: TIGR01393; LepA.1.  
 DR TIGRPFAM: TIGR00231; small\_GTP.1.  
 DR PROSITE: PS00301; EFACITOR\_GTP.1.  
 DR GTP-binding; Complete proteome.  
 KW NP\_BIND 51 58 GTP (BY SIMILARITY).  
 FT NP\_BIND 120 124 GTP (BY SIMILARITY).  
 FT NP\_BIND 174 177 GTP (BY SIMILARITY).  
 SO SEQUENCE 646 AA; 71329 MW; 256DC1AE894C4A3 CRC64;

Query Match 45.8%; Score 44; DB 1; Length 646;  
 Best Local Similarity 71.4%; Pred. No. 19;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGV 14  
 1 |||||1111  
 DB 57 KSTLADRMQLTGV 70

RESULT 12

LEPA\_MYCTU STANDARD: PRT: 653 AA.  
 AC P71739;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE GTP-binding protein lepa.  
 GN LEPA OR RV2404C OR WT2476 OR MTCY253.16.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Actinobacteria; Actinomycetales;  
 OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_Taxid-1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RX MEDLINE-98295987; PubMed-9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekla F.,  
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holtroyd S.,  
 Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 Sulton J.E., Taylor K., Whitehead S., Barrett B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey O.,  
 Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 laboratory strains.";  
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
 CC LEPA SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z81368; CAB03723.1; --  
 DR EMBL: AE007086; AAK46772.1; --  
 DR PIR: G70683; G70683.  
 DR HSSP: P13551; 2EFG.  
 DR TIGR: WT2476; --  
 DR Tuberculist: RV2404C; --  
 DR HAMAP: MF\_00071: -- 1.  
 DR InterPro: IPR000795; EF\_GTPbind.  
 DR InterPro: IPR000640; EFG\_C.  
 DR InterPro: IPR004161; EFTU\_D2.  
 DR InterPro: IPR006297; LepA.  
 DR InterPro: IPR005225; Small\_GTP.  
 DR Pfam: PF00679; EFG\_C.1.  
 DR Pfam: PF00009; GTP\_EFTU.1.  
 DR Pfam: PF03144; GTP\_EFTU\_D2.1.  
 DR PRINTS: PR00315; ELONGATNFCF.  
 DR TIGRPFAM: TIGR01393; LepA.1.  
 DR TIGRPFAM: TIGR00231; small\_GTP.1.  
 DR PROSITE: PS00301; EFACITOR\_GTP.1.  
 DR GTP-binding; Complete proteome.  
 KW NP\_BIND 59 66 GTP (BY SIMILARITY).  
 FT NP\_BIND 124 128 GTP (BY SIMILARITY).  
 FT NP\_BIND 178 181 GTP (BY SIMILARITY).  
 SO SEQUENCE 653 AA; 72395 MW; DA4AFB10E6C25755 CRC64;

Query Match 45.8%; Score 44; DB 1; Length 653;  
 Best Local Similarity 71.4%; Pred. No. 19;  
 Matches 10; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGV 14  
 1 |||||1111  
 DB 65 KSTLADRMQLTGV 78

RESULT 13

HS11\_CAEBL STANDARD: PRT: 145 AA.  
 ID HS11\_CAEBL  
 AC P34696;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Heat shock protein HSP16-1.  
 GN (HSP16-1A OR T27E4.2) AND (HSP16-1B OR T27E4.8).  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_Taxid-6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-85295957; PubMed-4033652;  
 RA Russnak R.H., Candido E.P.M.;  
 RT "Locus encoding a family of small heat shock genes in Caenorhabditis  
 RT elegans: two genes duplicated to form a 3.8-kilobase inverted  
 RT repeat.";  
 RL Mol. Cell. Biol. 5:1268-1278(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Bristol N2;  
 RA Bradehaw H.;  
 RT Submitted (Jul-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 41-63 FROM N.A.  
 RX MEDLINE-87231065; PubMed-3588308;  
 RA Kay R.J., Russnak R.H., Jones D., Mathias C., Candido E.P.M.;  
 RT "Expression of intron-containing C. elegans heat shock genes in mouse  
 RT cells demonstrates divergence of 3' splice site recognition sequences  
 RT between nematodes and vertebrates, and an inhibitory effect of heat  
 RT shock on the mammalian splicing apparatus.";

RL Nucleic Acids Res. 15:3723-3741(1987).  
 CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: K0373; AAA28068.1; -  
 CC EMBL: U64837; AAB04842.1; -  
 CC DR EMBL: U64837; AAB04839.1; -  
 CC DR PIR: B24289; B24289.  
 CC DR Wormpep: T27E4.2; CE14249.  
 CC DR Wormpep: T27E4.8; CE14249.  
 CC DR InterPro: IPR001436; Crystal1in\_alpha.  
 CC DR InterPro: IPR002068; HSP20.  
 CC PIRam: PF00011; HSP20; 1.  
 CC PRINTS: PRO0299; ACRYSTALLIN.  
 CC PROSITE: PS01031; HSP20; 1.  
 CC KW Heat shock; Multigene family.  
 CC SO SEQUENCE 145 AA; 16253 MW; 06C36A1F06D15A11 CRC64;  
 OY Query Match 44.8%; Score 43; DB 1; Length 145;  
 Best Local Similarity 45.0%; Pred. No. 5.9;  
 Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;  
 Db 1 KINADRLSLGSGVEIKQ 20  
 65 KINLGHSLTSGDEKTE 84  
 ID POP6\_YEAST STANDARD; PRT; 158 AA.  
 AC P53218;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ribonucleases P/MRP protein subunit POP6 (EC 3.1.26.5) (RNases P/MRP  
 DE 18.2 kDa subunit) (RNA processing protein POP6).  
 GN POP6 OR YGR030C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 CC NCBI\_TaxID=4932;  
 [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN-S288C;  
 CC MEDLINE=97435481; PubMed=9290212;  
 CC Rieger M., Brueckner M., Schaefer M., Mueller-Auer S.;  
 CC "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae  
 CC chromosome VII."  
 CC Yeast 13:1077-1090(1997).  
 CC -1- FUNCTION: COMPONENT OF RIBONUCLEASE P, A PROTEIN COMPLEX THAT  
 CC GENERATES MATURE tRNA MOLECULES BY CLEAVING THEIR 5' ENDS.  
 CC ALSO A COMPONENT OF RNASE MRP.  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of RNA, removing 5'-  
 CC extra-nucleotide from tRNA precursor.  
 CC -1- SUBUNIT: COMPONENT OF NUCLEAR RNASE P AND RNASE MRP RNASE P  
 CC RIBONUCLEOPROTEINS. RNASE P CONSISTS OF A RNA MOETY AND AT LEAST  
 CC 8 PROTEIN SUBUNITS; POP1, POP3, POP4, POP5, POP6, POP7, POP8 AND  
 CC RPL1.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (potential).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: Z72815; CA97018.1; -  
 CC DR PIR: S64321; S64321.  
 CC DR SGP: S0003262; POP6.  
 CC DR GO: GO:0000172; C:ribonuclease mitochondrial RNA processing c. . . ; IDA.  
 CC DR GO: GO:0005655; C:ribonuclease P complex; IDA.  
 CC DR GO: GO:0000171; F:ribonuclease MRP activity; IDA.  
 CC DR GO: GO:0004526; F:ribonuclease P activity; IDA.  
 CC KW Hydroxylase; Nuclear protein; tRNA processing; Coiled coil.  
 CC FT DOMAIN 51  
 CC SEQUENCE 158 AA; 18210 MW; 6C27A3FAD521181 CRC64;  
 OY Query Match 44.8%; Score 43; DB 1; Length 158;  
 Best Local Similarity 64.3%; Pred. No. 6.5;  
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Db 1 KINADRLSLGSGV 14  
 65 QINADRLSLGQGV 78  
 ID Y454\_YERPE STANDARD; PRT; 180 AA.  
 AC Q821P1;  
 DT 15-SEP-2003 (Rel. 42, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical UPF0244 protein YPO0454/Y3725.  
 GN YPO0454 OR Y3725.  
 OS Yersinia pestis.  
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 CC Enterobacteriaceae; Yersinia.  
 CC NCBI\_TaxID=632;  
 [1]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN-CO-92 / Biovar Orientalis;  
 CC MEDLINE=21470413; PubMed=11586360;  
 CC Parkhill J., Wren B.W., Thomson N.R., Tittball R.W., Holden M.T.G.,  
 CC Prentice M.B., Sebahinia M., James K.D., Churcher C., Mungall K.L.,  
 CC Baser S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,  
 CC Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 CC Fellwell T., Hamlin N., Holroyd S., Jagsels K., Karlyshev A.V.,  
 CC Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 CC Simmonds M., Skelton J., Stevens K., Whitehead S., Barrrell B.G.,  
 CC "Genome sequence of Yersinia pestis, the causative agent of plague."  
 CC Nature 413:523-527(2001).  
 [2]  
 CC SEQUENCE FROM N.A.  
 CC STRAIN-KIM5 / Biovar Mediaevalis;  
 CC MEDLINE=22137863; PubMed=12142430;  
 CC Deng W., Burland V., Plunkett G., III, Boutin A., Mayhew G.F., Liss P.,  
 CC Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
 CC Fetherston J.D., Lindler L.E., Brubaker G.V., Plano G.V.,  
 CC Straley R.D., McDonough K.A., Nilles M.L., Watson J.S., Blattner F.R.,  
 CC "Genome sequence of Yersinia pestis KIM."  
 CC J. Bacteriol. 184:4601-4611(2002).  
 CC -1- SIMILARITY: Belongs to the UPF0244 family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AJ14142; CAC9310.1; -  
 CC EMBL: AE013976; AAB67273.1; -  
 CC PIR: AC0056; AC0056.

DR HAMAP: MF\_00648; -: 1.  
DR InterPro: IPR002786; DUF84.  
DR Pfam: PF01931; DUF84; 1.  
DR ProDom: PD016591; DUF84; 1.  
DR TIGRFAMs: TIGR00258; TIGR00258; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 180 AA; 19694 MW; F95E9653D710828D CRC64;

Query Match 44.8%; Score 43; DB 1; Length 180;  
Best Local Similarity 47.1%; Pred. No. 7.4;  
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 4 LADRMGLSGVQDEIKQ 20  
| | | | | : | : |  
DB 120 LGDEMAVLGISNVRKQ 136

Search completed: October 1, 2003, 09:57:54  
Job time : 8.58824 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 33.1765 seconds  
(without alignments)  
155.564 Million cell updates/sec

Title: US-09-171-432a-46  
Perfect score: 96  
Sequence: 1 KINLADRMGLSGVQEIKEQ 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP\_ARCHAEA:\*  
2: SP\_BACTERIA:\*  
3: SP\_FUNGI:\*  
4: SP\_HUMAN:\*  
5: SP\_INVERTEBRATE:\*  
6: SP\_MAMMAL:\*  
7: SP\_MHC:\*  
8: SP\_ORGANELLE:\*  
9: SP\_PHAGE:\*  
10: SP\_PLANT:\*  
11: SP\_RODENT:\*  
12: SP\_VIRUS:\*  
13: SP\_VERTEBRATE:\*  
14: SP\_UNCLASSIFIED:\*  
15: SP\_RVIRUS:\*  
16: SP\_BACTERIAP:\*  
17: SP\_ARCHAEP:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	251	12 Q9ENP8	Q9enp8 hepatitis a
2	96	100.0	251	12 Q9ENN9	Q9enn9 hepatitis a
3	96	100.0	251	12 Q9ENQ4	Q9enq4 hepatitis a
4	96	100.0	251	12 Q9ENQ2	Q9enq2 hepatitis a
5	96	100.0	251	12 Q9ENN2	Q9enn2 hepatitis a
6	96	100.0	251	12 Q9ENP2	Q9enp2 hepatitis a
7	96	100.0	251	12 Q9ENQ6	Q9enq6 hepatitis a
8	96	100.0	251	12 Q9ENR1	Q9enr1 hepatitis a
9	96	100.0	251	12 Q9ENP1	Q9enp1 hepatitis a
10	96	100.0	251	12 Q9ENP4	Q9enp4 hepatitis a
11	96	100.0	251	12 Q9ENQ9	Q9enq9 hepatitis a
12	96	100.0	251	12 Q9ENN4	Q9enn4 hepatitis a
13	96	100.0	251	12 Q9ENP5	Q9enp5 hepatitis a
14	96	100.0	251	12 Q9ENQ3	Q9enq3 hepatitis a
15	96	100.0	251	12 Q9ENP7	Q9enp7 hepatitis a
16	96	100.0	251	12 Q9ENQ5	Q9enq5 hepatitis a

17	96	100.0	251	12 Q9ENP9	Q9enp9 hepatitis a
18	96	100.0	251	12 Q9ENN6	Q9enn6 hepatitis a
19	96	100.0	251	12 Q9ENP6	Q9enp6 hepatitis a
20	96	100.0	251	12 Q9ENQ8	Q9enq8 hepatitis a
21	96	100.0	251	12 Q9ENP3	Q9enp3 hepatitis a
22	96	100.0	251	12 Q9ENQ0	Q9enq0 hepatitis a
23	96	100.0	1124	12 Q84780	Q84780 hepatitis a
24	96	100.0	1161	12 Q05794	Q05794 hepatitis a
25	96	100.0	2216	12 Q9WMA2	Q9wma2 hepatitis a
26	96	100.0	2218	12 Q67824	Q67824 hepatitis a
27	96	100.0	2218	12 Q67817	Q67817 hepatitis a
28	96	100.0	2227	12 Q9DLJ2	Q9dlj2 hepatitis a
29	96	100.0	2227	12 Q9WMA0	Q9wma0 hepatitis a
30	96	100.0	2227	12 Q9WMA3	Q9wma3 hepatitis a
31	96	100.0	2227	12 Q67825	Q67825 hepatitis a
32	96	100.0	2227	12 Q9WV99	Q9wv99 hepatitis a
33	96	100.0	2227	12 Q8OV03	Q8ov03 hepatitis a
34	96	100.0	2227	12 Q9WMA1	Q9wma1 hepatitis a
35	96	100.0	2227	12 Q67826	Q67826 hepatitis a
36	96	100.0	2227	12 Q8VON6	Q8von6 hepatitis a
37	96	100.0	2227	12 Q9IFH5	Q9ifh5 hepatitis a
38	96	100.0	2227	12 Q9WMA4	Q9wma4 hepatitis a
39	96	100.0	251	12 Q9ENP0	Q9enp0 hepatitis a
40	96	100.0	251	12 Q9ENN5	Q9enn5 hepatitis a
41	96	100.0	251	12 Q9ENN7	Q9enn7 hepatitis a
42	96	100.0	2225	12 Q9DWR1	Q9dwr1 hepatitis a
43	96	100.0	2227	12 Q8OR16	Q8or16 hepatitis a
44	96	100.0	251	12 Q9ENR0	Q9enr0 hepatitis a
45	96	100.0	251	12 Q9ENN8	Q9enn8 hepatitis a

## ALIGNMENTS

RESULT 1  
ID Q9ENP8 PRELIMINARY; PRT; 251 AA.  
AC Q9ENP8;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A304;  
RA Fujiwara K.;  
RT "hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047665; BAB12173.1; -.  
FT NON\_TER  
FT NON\_TER  
SQ SEQUENCE 251 AA; 28693 MW; C32AD5651506751A CRC64;  
  
Query Match 100.0%; Score 96; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 KINLADRMGLSGVQEIKEQ 20  
DB 125 KINLADRMGLSGVQEIKEQ 144  
  
RESULT 2  
ID Q9ENN9 PRELIMINARY; PRT; 251 AA.  
AC Q9ENN9;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A68;  
RA Fujiwara K.;  
RT "Hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047674; BAB12182.1; -  
FT NON\_TER 1 1  
FT SEQUENCE 251 AA; 28666 MW; 5A315F46C830D214 CRC64;  
SQ

Query Match 100.0%; Score 96; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KINLADRLGLSGVQEIKEQ 20  
125 KINLADRLGLSGVQEIKEQ 144

RESULT 3  
Q9ENQ4 PRELIMINARY; PRT; 251 AA.  
ID Q9ENQ4;  
AC Q9ENQ4;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A201;  
RA Fujiwara K.;  
RT "Hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047659; BAB12167.1; -  
FT NON\_TER 1 1  
FT SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;  
SQ

Query Match 100.0%; Score 96; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KINLADRLGLSGVQEIKEQ 20  
125 KINLADRLGLSGVQEIKEQ 144

RESULT 4  
Q9ENQ2 PRELIMINARY; PRT; 251 AA.  
ID Q9ENQ2;  
AC Q9ENQ2;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A205;  
RA Fujiwara K.;

RT "Hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047661; BAB12169.1; -  
FT NON\_TER 1 1  
FT SEQUENCE 251 AA; 28723 MW; 558A2D3664C7343C CRC64;  
SQ

Query Match 100.0%; Score 96; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KINLADRLGLSGVQEIKEQ 20  
125 KINLADRLGLSGVQEIKEQ 144

RESULT 5  
Q9ENN2 PRELIMINARY; PRT; 251 AA.  
ID Q9ENN2;  
AC Q9ENN2;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A9;  
RA Fujiwara K.;  
RT "Hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047681; BAB12189.1; -  
FT NON\_TER 1 1  
FT SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;  
SQ

Query Match 100.0%; Score 96; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 KINLADRLGLSGVQEIKEQ 20  
125 KINLADRLGLSGVQEIKEQ 144

RESULT 6  
Q9ENP2 PRELIMINARY; PRT; 251 AA.  
ID Q9ENP2;  
AC Q9ENP2;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A5;  
RA Fujiwara K.;  
RT "Hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB047671; BAB12179.1; -  
FT NON\_TER 1 1  
FT SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;  
SQ

Query Match 100.0%; Score 96; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20  
|||||  
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 7  
O9EN06 PRELIMINARY; PRT; 251 AA.

AC O9EN06: 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;

(1)  
SEQUENCE FROM N.A.

RA Fujikura K.;  
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047657; BAB12165.1; -.

FT NON\_TER 1 1  
SQ SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6FD3B CRC64;

Query Match  
Best Local Similarity 100.0%; Score 96; DB 12; Length 251;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20  
|||||  
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 8  
O9ENR1 PRELIMINARY; PRT; 251 AA.

AC O9ENR1: 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;

(1)  
SEQUENCE FROM N.A.

RA Fujikura K.;  
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047652; BAB12160.1; -.

FT NON\_TER 1 1  
SQ SEQUENCE 251 AA; 28749 MW; 58A520D873893445 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 96; DB 12; Length 251;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20  
|||||  
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 9  
O9ENP1

ID O9ENP1 PRELIMINARY; PRT; 251 AA.

AC O9ENP1: 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;

(1)  
SEQUENCE FROM N.A.

RA Fujikura K.;  
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047672; BAB12180.1; -.

FT NON\_TER 1 1  
SQ SEQUENCE 251 AA; 28663 MW; C7EA66FBD19A1619 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 96; DB 12; Length 251;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20  
|||||  
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 10  
O9ENP4 PRELIMINARY; PRT; 251 AA.

AC O9ENP4: 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;

(1)  
SEQUENCE FROM N.A.

RA Fujikura K.;  
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047669; BAB12177.1; -.

FT NON\_TER 1 1  
SQ SEQUENCE 251 AA; 28666 MW; 5A315F46C830D214 CRC64;

Query Match  
Best Local Similarity 100.0%; Score 96; DB 12; Length 251;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20  
|||||  
DB 125 KINLADRMGLSGVOEIKQ 144

RESULT 11  
O9EN09 PRELIMINARY; PRT; 251 AA.

AC O9EN09: 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.

OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A159;  
RA Fujiwara K.;  
RT "hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047654; BAB12162.1; -.  
FT NON\_TER 1  
FT NON\_TER 251  
SQ SEQUENCE 251 AA; 28720 MM; C3342482882F19CA CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DY 1 KINLADRMGLSGVOEIKQ 20  
Db 125 KINLADRMGLSGVOEIKQ 144

RESULT 12  
OGENN4 PRELIMINARY; PRT; 251 AA.  
ID OGENN4  
AC OGENN4;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A77;  
RA Fujiwara K.;  
RT "hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047679; BAB12187.1; -.  
FT NON\_TER 1  
FT NON\_TER 251  
SQ SEQUENCE 251 AA; 28734 MM; 3B96B782882F19D9 CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KINLADRMGLSGVOEIKQ 20  
125 KINLADRMGLSGVOEIKQ 144

RESULT 13  
OGENP5 PRELIMINARY; PRT; 251 AA.  
ID OGENP5  
AC OGENP5;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AA07;  
RA Fujiwara K.;  
RT "hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047668; BAB12176.1; -.  
FT NON\_TER 1  
FT NON\_TER 1

FT NON\_TER 251  
SQ SEQUENCE 251 AA; 28614 MM; 8334EF19C757A6D CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20  
Db 125 KINLADRMGLSGVOEIKQ 144

RESULT 14  
OGENO3 PRELIMINARY; PRT; 251 AA.  
ID OGENO3  
AC OGENO3;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A204;  
RA Fujiwara K.;  
RT "hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047660; BAB12168.1; -.  
FT NON\_TER 1  
FT NON\_TER 251  
SQ SEQUENCE 251 AA; 28693 MM; C32AD5651506751A CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20  
Db 125 KINLADRMGLSGVOEIKQ 144

RESULT 15  
OGENP7 PRELIMINARY; PRT; 251 AA.  
ID OGENP7  
AC OGENP7;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A306;  
RA Fujiwara K.;  
RT "hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047666; BAB12174.1; -.  
FT NON\_TER 1  
FT NON\_TER 251  
SQ SEQUENCE 251 AA; 28720 MM; C3342482882F19CA CRC64;

Query Match 100.0%; Score 96; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.6e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKQ 20  
1 KINLADRMGLSGVOEIKQ 20

Db 125 KINLADRMIGSGVOEIKEQ 144

Search completed: October 1, 2003, 10:02:41  
Job time : 34.1765 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 ; Search time 12.7059 Seconds  
(without alignments)  
66.600 Million cell updates/sec

Title: US-09-171-432a-46  
Perfect score: 96  
Sequence: 1 KINLADRMGLSGVOEIKEQ 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/Backfilist1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	96	100.0	2227	3	US-08-475-886-2
2	96	100.0	2227	3	US-08-475-886-4
3	96	100.0	2227	3	US-08-475-886-6
4	96	100.0	2227	3	US-08-397-232-2
5	96	100.0	2227	3	US-08-397-232-4
6	96	100.0	2227	3	US-09-171-387-2
7	96	100.0	2227	4	US-09-653-499-2
8	96	100.0	2227	4	US-09-653-499-4
9	96	100.0	2227	4	US-09-653-499-6
10	96	100.0	2227	4	US-09-653-499-8
11	42	43.8	1250	1	US-08-441-139-9
12	41	42.7	366	3	US-08-746-883-6
13	41	42.7	677	4	US-09-252-991A-32924
14	41	42.7	708	4	US-09-643-597-369
15	40.5	42.2	295	4	US-09-328-352-5192
16	40	41.7	331	4	US-09-634-238-247
17	40	41.7	332	1	US-08-469-649-2
18	40	41.7	332	4	US-09-347-878-60
19	40	41.7	640	2	US-09-198-452A-44
20	40	41.7	643	3	US-08-846-762-95
21	40	41.7	186	3	US-09-178-252-25
22	40	41.7	1207	1	US-07-951-715A-7
23	40	41.7	1207	2	US-08-459-448A-7
24	40	41.7	1207	3	US-08-459-595A-7
25	40	41.7	1207	3	US-08-459-504B-7
26	40	41.7	1207	3	US-08-459-44A-7
27	40	41.7	1207	3	US-09-053-549-8

28	40	41.7	1207	4	US-09-547-422-7	Sequence 7, App11
29	40	41.7	1227	1	US-08-448-170-8	Sequence 8, App11
30	40	41.7	1227	3	US-09-053-549-2	Sequence 2, App11
31	40	41.7	1227	3	US-08-961-803-9	Sequence 9, App11
32	40	41.7	1229	1	US-08-100-709-4	Sequence 4, App11
33	40	41.7	1229	1	US-08-176-865-4	Sequence 4, App11
34	40	41.7	1229	1	US-08-474-038-4	Sequence 4, App11
35	40	41.7	1229	2	US-08-779-046-4	Sequence 4, App11
36	40	41.7	1297	2	US-09-107-532A-4552	Sequence 4, App11
37	40	41.7	1297	2	US-08-685-576-1	Sequence 4, App11
38	40	41.7	1388	2	US-09-724-623-122	Sequence 1, App11
39	39	40.6	310	4	US-09-134-001C-4689	Sequence 122, App
40	39	40.6	318	4	US-09-252-991A-20350	Sequence 4689, Ap
41	39	40.6	528	3	US-08-904-871-5	Sequence 20350, A
42	39	40.6	593	4	US-09-252-991A-25274	Sequence 5, App11
43	39	40.6	748	3	US-08-904-871-6	Sequence 25274, A
44	39	40.6	748	3	US-08-904-871-13	Sequence 6, App11
45	39	40.6	748	3	US-08-904-871-13	Sequence 13, App11

## ALIGNMENTS

```
RESULT 1
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2

Query Match          100.0%; Score 96; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 KINLADRMGLSGVOEIKEQ 20
Db      961 KINLADRMGLSGVOEIKEQ 980

RESULT 2
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
```

;; EARLIER FILING DATE: 1995-03-10  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 4  
;; LENGTH: 2227  
;; TYPE: PRT  
;; ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
US-08-475-886-4

Query Match 100.0%; Score 96; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRLGLSGVQEIKEQ 20  
Db 961 KINLADRLGLSGVQEIKEQ 980

## RESULT 3

US-08-475-886-6  
Sequence 6, Application US/08475866A  
Patent No. 6113912

;; GENERAL INFORMATION:  
;; APPLICANT: FUNKHOUSER, ANN W  
;; APPLICANT: EMERSON, SUZANNE U  
;; APPLICANT: PURCELL, ROBERT H  
;; APPLICANT: D'HONDT, ERIC  
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
;; FILE REFERENCE: 20264262US2  
;; CURRENT FILING DATE: 1995-06-07  
;; EARLIER APPLICATION NUMBER: 07/947,338  
;; EARLIER FILING DATE: 1992-09-18  
;; EARLIER APPLICATION NUMBER: 08/397,232  
;; EARLIER FILING DATE: 1995-03-10  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 6  
;; LENGTH: 2227  
;; TYPE: PRT  
;; ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-08-475-886-6

Query Match 100.0%; Score 96; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRLGLSGVQEIKEQ 20  
Db 961 KINLADRLGLSGVQEIKEQ 980

## RESULT 4

US-08-397-232-2  
Sequence 2, Application US/08397232A  
Patent No. 6180110

;; GENERAL INFORMATION:  
;; APPLICANT: FUNKHOUSER, ANN W  
;; APPLICANT: EMERSON, SUZANNE U  
;; APPLICANT: PURCELL, ROBERT H  
;; APPLICANT: D'HONDT, ERIC  
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
;; FILE REFERENCE: 20264262US1  
;; CURRENT APPLICATION NUMBER: US/08/397,232A  
;; EARLIER FILING DATE: 1995-04-17  
;; EARLIER APPLICATION NUMBER: 07/947,338  
;; EARLIER FILING DATE: 1992-09-18  
;; EARLIER APPLICATION NUMBER: PCT/US93/08610  
;; NUMBER OF SEQ ID NOS: 4  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 2  
;; LENGTH: 2227

;; TYPE: PRT  
;; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-08-397-232-2

Query Match 100.0%; Score 96; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRLGLSGVQEIKEQ 20  
Db 961 KINLADRLGLSGVQEIKEQ 980

## RESULT 5

US-08-397-232-4  
Sequence 4, Application US/08397232A  
Patent No. 6180110

;; GENERAL INFORMATION:  
;; APPLICANT: FUNKHOUSER, ANN W  
;; APPLICANT: EMERSON, SUZANNE U  
;; APPLICANT: PURCELL, ROBERT H  
;; APPLICANT: D'HONDT, ERIC  
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
;; FILE REFERENCE: 20264262US1  
;; CURRENT APPLICATION NUMBER: US/08/397,232A  
;; CURRENT FILING DATE: 1995-04-17  
;; EARLIER FILING DATE: 1992-09-18  
;; EARLIER APPLICATION NUMBER: 07/947,338  
;; EARLIER FILING DATE: 1993-09-17  
;; EARLIER APPLICATION NUMBER: PCT/US93/08610  
;; NUMBER OF SEQ ID NOS: 4  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 4  
;; LENGTH: 2227  
;; TYPE: PRT  
;; ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-08-397-232-4

Query Match 100.0%; Score 96; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRLGLSGVQEIKEQ 20  
Db 961 KINLADRLGLSGVQEIKEQ 980

## RESULT 6

US-09-171-387-2  
Sequence 2, Application US/09171387  
Patent No. 6280734

;; GENERAL INFORMATION:  
;; APPLICANT: RAYCHAUDHURI, GOPA;  
;; EMERSON, SUZANNE, U.;  
;; PURCELL, ROBERT, H.;  
;; TITLE OF INVENTION: SIMIAN-HUMAN HAV  
;; HAVING A CHIMERIC 2C PROTEIN  
;; NUMBER OF SEQUENCES: 3  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESS: MORGAN & FINNEGAN, L.L.P.  
;; STREET: 345 PARK AVENUE  
;; CITY: NEW YORK  
;; STATE: NEW YORK  
;; COUNTRY: USA  
;; ZIP: 10154  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: FLOPPY DISK  
;; COMPUTER: IBM PC COMPATIBLE  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: MICROSOFT WORD 97  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/171,387  
;; FILING DATE: 24-Mar-1999



PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US97/06506  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US60/015,642  
FILING DATE: 19-APR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: William S. Feller  
REGISTRATION NUMBER: 26,728  
REFERENCE/DOCKET NUMBER: 2026-4229051  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 751-6849  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 2  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2227 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-171-387-2

Query Match 100.0%; Score 96; DB 4; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVGEIRKQ 20  
DB 961 KINLADRMGLSGVGEIRKQ 980

RESULT 7  
US-09-653-499-2  
Sequence 2, Application US/09653499  
Patent No. 6423318  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262US2  
CURRENT APPLICATION NUMBER: US/09/653,499  
CURRENT FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: 08/475,886  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-09-653-499-2

Query Match 100.0%; Score 96; DB 4; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVGEIRKQ 20  
DB 961 KINLADRMGLSGVGEIRKQ 980

RESULT 8  
US-09-653-499-4  
Sequence 4, Application US/09653499  
Patent No. 6423318  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U

APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262US2  
CURRENT APPLICATION NUMBER: US/09/653,499  
CURRENT FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: 08/475,886  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
US-09-653-499-4

Query Match 100.0%; Score 96; DB 4; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVGEIRKQ 20  
DB 961 KINLADRMGLSGVGEIRKQ 980

RESULT 9  
US-09-653-499-6  
Sequence 6, Application US/09653499  
Patent No. 6423318  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262US2  
CURRENT APPLICATION NUMBER: US/09/653,499  
CURRENT FILING DATE: 2000-08-31  
PRIOR APPLICATION NUMBER: 08/475,886  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-09-653-499-6

Query Match 100.0%; Score 96; DB 4; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.8e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVGEIRKQ 20  
DB 961 KINLADRMGLSGVGEIRKQ 980

RESULT 10  
US-08-441-139-9  
Sequence 9, Application US/08441139  
Patent No. 5773245  
GENERAL INFORMATION:  
APPLICANT: Witterup, Dr. Karl D.  
APPLICANT: Robinson, Anne S.  
TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF

TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: NY  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/441,139  
FILING DATE: 15-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/089,997  
FILING DATE: 06-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: DIGILIO, Frank S.  
REGISTRATION NUMBER: 31,346  
TELEPHONE: 516-742-4343  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1250 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-441-139-9

Query Match 43.8%; Score 42; DB 1; Length 1250;  
Best Local Similarity 42.1%; Pred. NO. 57;  
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 2 INLADRMGLSGVOEIKR 20  
DB 871 INLIDRLGFGVINEIKR 889

RESULT 11  
US-08-746-883-6  
Sequence 6, Application US/08746883  
Patent No. 6093693

GENERAL INFORMATION:  
APPLICANT: Julius, Michael H., Philipp, Dominik,  
APPLICANT: Alizadeh-Khivari, Kamel  
TITLE OF INVENTION: B Cell Activation  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Blake, Cassels & Graydon  
STREET: Box 25, Commerce Court West  
CITY: Toronto  
COUNTRY: Canada  
ZIP: M5L 1A9  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3 1/2 inch, 1.4 Mb storage  
COMPUTER: COMPAQ, IBM PC compatible  
OPERATING SYSTEM: MS-DOS 5.1  
SOFTWARE: WORD PERFECT  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/746,883  
FILING DATE: NO. 6093693ember 18, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunt, John C.  
REGISTRATION NUMBER: 36,424  
REFERENCE/DOCKET NUMBER: 47841/00008

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 863-4344  
TELEFAX: (416) 863-2653  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 366 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-746-883-6

Query Match 42.7%; Score 41; DB 3; Length 366;  
Best Local Similarity 63.6%; Pred. NO. 19;  
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

OY 7 RMLGSGVOEI 17  
DB 107 RVLGISGLQEL 117

RESULT 12  
US-09-252-991A-32924  
Sequence 32924, Application US/09252991A  
Patent No. 6551795

GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 32924  
LENGTH: 677  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-32924

Query Match 42.7%; Score 41; DB 4; Length 677;  
Best Local Similarity 47.4%; Pred. NO. 41;  
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 KINLADRMGLSGVOEIKR 19  
DB 90 KTLVDKILKSLTIDRKE 108

RESULT 13  
US-09-643-597-369  
Sequence 369, Application US/09643597  
Patent No. 6426072

GENERAL INFORMATION:  
APPLICANT: Wang, Tongtong  
APPLICANT: Fan, Liqun  
APPLICANT: Kalos, Michael D.  
APPLICANT: Bangur, Chaitanya S.  
APPLICANT: Hosken, Nancy  
APPLICANT: Fanger, Gary R.  
APPLICANT: Li, Samuel X.  
APPLICANT: Wang, Aljun  
APPLICANT: Skelky, Yasir A.W.  
APPLICANT: Henderson, Robert A.  
APPLICANT: McNeill, Patricia D.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
FILE REFERENCE: 210121.455C11  
CURRENT APPLICATION NUMBER: US/09/643,597  
CURRENT FILING DATE: 2000-08-21  
NUMBER OF SEQ ID NOS: 369

SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 369  
LENGTH: 708  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-643-597-369

Query Match 42.7%; Score 41; DB 4; Length 708;  
Best Local Similarity 38.9%; Pred. No. 43;  
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 3 NLADRMGLSGVOEIRKQ 20  
Db 373 NHKRSIGTGIGIDLER 390

RESULT 14  
US-09-328-352-5192  
Sequence 5192, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Breton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 5192  
LENGTH: 295  
TYPE: PRT  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-5192

Query Match 42.2%; Score 40.5; DB 4; Length 295;  
Best Local Similarity 47.6%; Pred. No. 18;  
Matches 10; Conservative 4; Mismatches 4; Indels 3; Gaps 1;

OY 3 NLADRM--LGLSGVOEIRKQ 20  
Db 267 NLADKVALNTSISTIOEIRKQ 287

RESULT 15  
US-09-634-238-247  
Sequence 247, Application US/09634238  
Patent No. 6544772  
GENERAL INFORMATION:  
APPLICANT: Glenn, Matthew  
APPLICANT: Havukkala, Ilkka J.  
APPLICANT: Bloksberg, Leonard, N.  
APPLICANT: Lubbers, Mark W.  
APPLICANT: Dekker, James  
APPLICANT: Christenson, Anna C.  
APPLICANT: Holland, Ross  
APPLICANT: O'Toole, Paul W.  
APPLICANT: Reid, Julian R.  
APPLICANT: Coolbear, Timothy  
TITLE OF INVENTION: Polynucleotides, materials incorporating  
FILE REFERENCE: 11000, 1043U1  
CURRENT APPLICATION NUMBER: US/09/634,238  
CURRENT FILING DATE: 2000-08-08  
NUMBER OF SEQ ID NOS: 422  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 247  
LENGTH: 331  
TYPE: PRT  
ORGANISM: Lactobacillus rhamnosus  
US-09-634-238-247

Query Match 41.7%; Score 40; DB 4; Length 331;  
Best Local Similarity 42.1%; Pred. No. 26;

Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;  
OY 2 INLADRMGLSGVOEIRKQ 20  
Db 259 ISFLNQMONVDGYOIRKQ 277

Search completed: October 1, 2003, 10:06:33  
Job time: 13.7059 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 ; Search time 24.1176 Seconds  
(without alignments)  
131.201 Million cell updates/sec

Title: US-09-171-432a-46

Perfect score: 96

Sequence: 1 KINLADRMGLSGVDEIKQ 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Minimum number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications, AA.\*  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubppaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	96	100.0	352	15	US-10-272-459-45
2	96	100.0	980	10	US-10-272-459-41
3	96	100.0	2227	10	US-09-929-955-12
4	96	100.0	2227	14	US-10-104-966-12
5	96	100.0	2227	14	US-10-135-988-2
6	96	100.0	2227	14	US-10-135-988-4
7	96	100.0	2227	14	US-10-135-988-6
8	45	46.9	632	15	US-10-156-761-13093
9	42	43.8	1250	10	US-09-801-368-364
10	41.5	43.2	1967	15	US-10-132-382-6
11	41.5	43.2	2013	15	US-10-132-382-2
12	41.5	43.2	2014	15	US-10-132-382-8
13	41.5	43.2	2040	15	US-10-132-382-4
14	41.5	42.7	366	10	US-09-870-759-72
15	41	42.7	366	12	US-09-751-708A-72

16	41	42.7	366	15	US-10-281-478-10	Sequence 10, Appl
17	41	42.7	605	9	US-09-815-242-12087	Sequence 12087, A
18	41	42.7	708	9	US-09-735-705-369	Sequence 369, App
19	41	42.7	708	10	US-09-850-716A-369	Sequence 369, App
20	41	42.7	708	10	US-09-897-778-369	Sequence 82, Appl
21	41	42.7	708	12	US-10-117-982-369	Sequence 369, App
22	41	42.7	1309	10	US-09-862-027-82	Sequence 369, App
23	40	41.7	140	10	US-09-862-027-82	Sequence 4301, Ap
24	40	41.7	154	9	US-09-864-761-34931	Sequence 34931, A
25	40	41.7	316	12	US-09-907-218-76	Sequence 76, Appl
26	40	41.7	549	9	US-09-815-242-10680	Sequence 10680, A
27	40	41.7	615	10	US-09-738-626-6067	Sequence 6067, Ap
28	40	41.7	638	15	US-10-216-209-95	Sequence 95, Appl
29	40	41.7	643	9	US-09-826-660-25	Sequence 25, Appl
30	40	41.7	649	14	US-10-023-437-27	Sequence 27, Appl
31	40	41.7	1151	10	US-09-749-601A-10	Sequence 10, Appl
32	40	41.7	1151	11	US-09-912-697-33	Sequence 33, Appl
33	40	41.7	1186	9	US-09-826-660-23	Sequence 23, Appl
34	40	41.7	1207	11	US-09-988-662-7	Sequence 7, Appl
35	40	41.7	1770	10	US-09-801-368-298	Sequence 298, Appl
36	39	40.6	110	9	US-09-864-761-48195	Sequence 48195, A
37	39	40.6	270	12	US-10-237-496-88	Sequence 88, Appl
38	39	40.6	270	12	US-10-242-074-88	Sequence 88, Appl
39	39	40.6	270	12	US-10-242-505-88	Sequence 88, Appl
40	39	40.6	270	12	US-10-242-574-88	Sequence 88, Appl
41	39	40.6	270	12	US-10-243-261-88	Sequence 88, Appl
42	39	40.6	270	12	US-10-243-282-88	Sequence 88, Appl
43	39	40.6	270	12	US-10-243-402-88	Sequence 88, Appl
44	39	40.6	270	12	US-10-243-431-88	Sequence 88, Appl
45	39	40.6	270	12	US-10-243-164-88	Sequence 88, Appl

## ALIGNMENTS

RESULT 1  
US-10-272-459-45  
; Sequence 45, Application US/10272459  
; Publication No. US20030124517A1  
; GENERAL INFORMATION:  
; APPLICANT: PICHUNTES, Sergio  
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT  
; FILE REFERENCE: PPI/955.002 / 2301-17955  
; CURRENT APPLICATION NUMBER: US/10/272,459  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 45  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein  
; OTHER INFORMATION: of 38.8 kda  
US-10-272-459-45

Query Match 100.0%; Score 96; DB 15; Length 352;  
Best Local Similarity 100.0%; Pred. No. 2.2e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KINLADRMGLSGVDEIKQ 20  
|||||  
Db 333 KINLADRMGLSGVDEIKQ 352

RESULT 2  
US-10-272-459-41  
; Sequence 41, Application US/10272459  
; Publication No. US20030124517A1  
; GENERAL INFORMATION:  
; APPLICANT: PICHUNTES, Sergio  
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

;; TITLE OF INVENTION: PROTEINS AND USES THEREOF  
;; FILE REFERENCE: P17955.002 / 2301-17955  
;; CURRENT APPLICATION NUMBER: US/10/272.459  
;; CURRENT FILING DATE: 2002-10-15  
;; NUMBER OF SEQ ID NOS: 50  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 41  
;; LENGTH: 980  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein  
;; OTHER INFORMATION: of 115.5 kda  
US-10-272-459-41

Query Match 100.0%; Score 96; DB 15; Length 980;  
Best Local Similarity 100.0%; Pred. No. 7.4e-08;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRLGLSGVOEIKQ 20  
|||||  
961 KINLADRLGLSGVOEIKQ 980

RESULT 3  
US-09-929-955-12  
;; Sequence 12, Application US/09929955  
;; Patent No. US20020136740A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Catharina Hultgren  
;; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND  
;; FILE REFERENCE: METHODS OF USE THEREOF  
;; CURRENT APPLICATION NUMBER: US/09/929.955  
;; CURRENT FILING DATE: 2001-08-15  
;; PRIOR APPLICATION NUMBER: 09/705.547  
;; PRIOR FILING DATE: 2000-11-03  
;; PRIOR APPLICATION NUMBER: 60/229.175  
;; PRIOR FILING DATE: 2000-08-29  
;; PRIOR APPLICATION NUMBER: 60/225.767  
;; PRIOR FILING DATE: 2000-08-17  
;; NUMBER OF SEQ ID NOS: 49  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 12  
;; LENGTH: 2227  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Hepatitis A virus sequence  
US-09-929-955-12

Query Match 100.0%; Score 96; DB 10; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRLGLSGVOEIKQ 20  
|||||  
Db 961 KINLADRLGLSGVOEIKQ 980

RESULT 4  
US-10-104-966-12  
;; Sequence 12, Application US/10104966  
;; Publication No. US20020155124A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Matti Salberg  
;; APPLICANT: Catharina Hultgren  
;; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND  
;; FILE REFERENCE: METHODS OF USE THEREOF  
;; CURRENT APPLICATION NUMBER: US/10/104.966  
;; CURRENT FILING DATE: 2002-03-22

;; PRIOR APPLICATION NUMBER: 09/705.547  
;; PRIOR FILING DATE: 2000-11-03  
;; PRIOR APPLICATION NUMBER: 60/229.175  
;; PRIOR FILING DATE: 2000-08-29  
;; NUMBER OF SEQ ID NOS: 15  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 12  
;; LENGTH: 2227  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Hepatitis A virus sequence  
US-10-104-966-12

Query Match 100.0%; Score 96; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRLGLSGVOEIKQ 20  
|||||  
Db 961 KINLADRLGLSGVOEIKQ 980

RESULT 5  
US-10-135-988-2  
;; Sequence 2, Application US/10135988  
;; Publication No. US20020176869A1  
;; GENERAL INFORMATION:  
;; APPLICANT: FUNKHOUSER, ANN W  
;; APPLICANT: EMERSON, SUZANNE U  
;; APPLICANT: PURCELL, ROBERT H  
;; APPLICANT: D'HONDT, ERIC  
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
;; FILE REFERENCE: 20264262US3  
;; CURRENT APPLICATION NUMBER: US/10/135.988  
;; CURRENT FILING DATE: 2002-04-29  
;; PRIOR APPLICATION NUMBER: 07/947.338  
;; PRIOR FILING DATE: 1992-09-18  
;; PRIOR APPLICATION NUMBER: 08/397.232  
;; PRIOR FILING DATE: 1995-03-10  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: PatentIn Ver. 2.1  
;; SEQ ID NO 2  
;; LENGTH: 2227  
;; TYPE: PRT  
;; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-10-135-988-2

Query Match 100.0%; Score 96; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KINLADRLGLSGVOEIKQ 20  
|||||  
Db 961 KINLADRLGLSGVOEIKQ 980

RESULT 6  
US-10-135-988-4  
;; Sequence 4, Application US/10135988  
;; Publication No. US20020176869A1  
;; GENERAL INFORMATION:  
;; APPLICANT: FUNKHOUSER, ANN W  
;; APPLICANT: EMERSON, SUZANNE U  
;; APPLICANT: PURCELL, ROBERT H  
;; APPLICANT: D'HONDT, ERIC  
;; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
;; FILE REFERENCE: 20264262US3  
;; CURRENT APPLICATION NUMBER: US/10/135.988  
;; CURRENT FILING DATE: 2002-04-29  
;; PRIOR APPLICATION NUMBER: 07/947.338  
;; PRIOR FILING DATE: 1992-09-18  
;; PRIOR APPLICATION NUMBER: 08/397.232

;; PRIOR FILING DATE: 1995-03-10  
;; NUMBER OF SEQ ID NOS: 6  
;; SOFTWARE: Patentln Ver. 2.1  
;; SEQ ID NO 4  
;; LENGTH: 2227  
;; TYPE: PRT  
;; ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
US-10-135-988-4

Query Match 100.0%; Score 96; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KINLADRMGLSGVQEIKEQ 20  
Db 961 KINLADRMGLSGVQEIKEQ 980

RESULT 7  
US-10-135-988-6  
Sequence 6, Application US/10135988  
Publication No. US20020176869A1  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN M  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262US3  
CURRENT APPLICATION NUMBER: US/10/135, 988  
CURRENT FILING DATE: 2002-04-29  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 6  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-10-135-988-6

Query Match 100.0%; Score 96; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2e-07;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KINLADRMGLSGVQEIKEQ 20  
961 KINLADRMGLSGVQEIKEQ 980

RESULT 8  
US-10-156-761-13093  
Sequence 13093, Application US/10156761  
Publication No. US20030119018A1  
GENERAL INFORMATION:  
APPLICANT: OMURA, SATOSHI  
APPLICANT: IKEDA, HARUO  
APPLICANT: ISHIKAWA, JUN  
APPLICANT: HORIKAWA, HIROSHI  
APPLICANT: SHIBA, TADAYOSHI  
APPLICANT: SAKAKI, YOSHIYUKI  
APPLICANT: HATTORI, MASAHIRA  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-262  
CURRENT APPLICATION NUMBER: US/10/156,761  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: JP 2001-204089  
PRIOR FILING DATE: 2001-05-30  
PRIOR APPLICATION NUMBER: JP 2001-272697  
PRIOR FILING DATE: 2001-08-02  
NUMBER OF SEQ ID NOS: 15109

;; SEQ ID NO 13093  
;; LENGTH: 622  
;; TYPE: PRT  
;; ORGANISM: Streptomyces avermitilis  
US-10-156-761-13093

Query Match 46.9%; Score 45; DB 15; Length 622;  
Best Local Similarity 52.6%; Pred. No. 27;  
Matches 10; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Oy 1 KINLADRMGLSGVQEIKE 19  
Db 32 KSTLADRMQLTGVVDORQ 50

RESULT 9  
US-09-801-368-364  
Sequence 364, Application US/09801368  
Patent No. US20020128250A1  
GENERAL INFORMATION:  
APPLICANT: Busby, Robert  
APPLICANT: Call, Brian  
APPLICANT: Hecht, Peter  
APPLICANT: Holtzman, Doug  
APPLICANT: Madden, Kevin  
APPLICANT: Maxon, Mary  
APPLICANT: Milne, Todd  
APPLICANT: No. US20020128250A1man, Thea  
APPLICANT: Royer, John  
APPLICANT: Salama, Sofie  
APPLICANT: Sherman, Amlr  
APPLICANT: Silva, Jeff  
APPLICANT: Summers, Eric  
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi  
FILE REFERENCE: 109272.147  
CURRENT APPLICATION NUMBER: US/09/801,368  
CURRENT FILING DATE: 2001-03-07  
PRIOR APPLICATION NUMBER: US 09/487,558  
PRIOR FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: US 60/160,587  
PRIOR FILING DATE: 1999-10-20  
NUMBER OF SEQ ID NOS: 440  
SOFTWARE: Patentln version 3.0  
SEQ ID NO 364  
LENGTH: 1250  
TYPE: PRT  
ORGANISM: Saccharomyces cerevisiae  
US-09-801-368-364

Query Match 43.8%; Score 42; DB 10; Length 1250;  
Best Local Similarity 42.1%; Pred. No. 2e+02;  
Matches 8; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Oy 2 INLADRMGLSGVQEIKEQ 20  
Db 871 UNLDRTLGFWVINEIKRK 889

RESULT 10  
US-10-132-382-6  
Sequence 6, Application US/10132382  
Publication No. US20030045699A1  
GENERAL INFORMATION:  
APPLICANT: WEISS, BERTRAM  
TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS  
FILE REFERENCE: SCH-1811  
CURRENT APPLICATION NUMBER: US/10/132,382  
CURRENT FILING DATE: 2002-04-26  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 6  
LENGTH: 1987  
TYPE: PRT

```

; ORGANISM: Homo sapiens
US-10-132-382-6

Query Match          43.2%  Score 41.5; DB 15; Length 1987;
Best Local Similarity 55.0%  Pred. No. 4.3e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1

      1 KINLADRMIG-LSGVOEIKE 19
      ||:| | | | | | |
Db      279 KNLSDNHLGELPGVQSSDE 298

RESULT 11
US-10-132-382-2
; Sequence 2, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2013
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-382-2

Query Match          43.2%  Score 41.5; DB 15; Length 2013;
Best Local Similarity 55.0%  Pred. No. 4.3e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1

      1 KINLADRMIG-LSGVOEIKE 19
      ||:| | | | | | |
Db      305 KNLSDNHLGELPGVQSSDE 324

RESULT 12
US-10-132-382-8
; Sequence 8, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS
; FILE REFERENCE: SCH-1811
; CURRENT APPLICATION NUMBER: US/10/132,382
; CURRENT FILING DATE: 2002-04-26
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 2014
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-132-382-8

Query Match          43.2%  Score 41.5; DB 15; Length 2014;
Best Local Similarity 55.0%  Pred. No. 4.3e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1

      1 KINLADRMIG-LSGVOEIKE 19
      ||:| | | | | | |
Db      306 KNLSDNHLGELPGVQSSDE 325

RESULT 13
US-10-132-382-4
; Sequence 4, Application US/10132382
; Publication No. US20030045699A1
; GENERAL INFORMATION:
; APPLICANT: WEISS, BERTRAM
; TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROIDS

```

```

FILE REFERENCE: SCH-1811
CURRENT APPLICATION NUMBER: US/10/132,382
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 26
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2040
TYPE: PRT
ORGANISM: Homo sapiens
US-10-132-382-4

Query Match                      43.2%; Score 41.5; DB 15; Length 2040;
Best Local Similarity 55.0%; Pred. No. 4,4e+02;
Matches 11; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY      1 KINIDRMIG-LSGVOEIKE 19
       |::|:| 11 1 11 1
Db      332 KLNIDNHGELPGVQSDE 351

RESULT 14
US-09-870-759-72
Sequence 72, Application US/09870759
Patent No. US20020177551A1
GENERAL INFORMATION:
APPLICANT: TERMAN, David S
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
FILE REFERENCE: 870759
CURRENT APPLICATION NUMBER: US/09/870,759
CURRENT FILING DATE: 2002-01-14
PRIORITY APPLICATION NUMBER: US 60/208,128
PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 166
SOFTWARE: PatentIn version 3.1
SEQ ID NO 72
LENGTH: 366
TYPE: PRT
ORGANISM: Mus musculus
US-09-870-759-72

Query Match                      42.7%; Score 41; DB 10; Length 366;
Best Local Similarity 63.6%; Pred. No. 70;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY      7 RMLGSVCDEI 17
       |::|:|:|:|
Db      107 RVLGISGLDEL 117

RESULT 15
US-09-751-708A-72
Sequence 72, Application US/09751708A
Publication No. US20030157113A1
GENERAL INFORMATION:
APPLICANT: TERMAN, David S
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
FILE REFERENCE: 751708
CURRENT APPLICATION NUMBER: US/09/751,708A
CURRENT FILING DATE: 2002-10-15
PRIORITY APPLICATION NUMBER: US 60/173,371
PRIOR FILING DATE: 1999-12-28
NUMBER OF SEQ ID NOS: 166
SOFTWARE: PatentIn version 3.1
SEQ ID NO 72
LENGTH: 366
TYPE: PRT
ORGANISM: Mus musculus
US-09-751-708A-72

Query Match                      42.7%; Score 41; DB 12; Length 366;
Best Local Similarity 63.6%; Pred. No. 70;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```



Fri Oct 2 09:14:00 2003

OY	7	RMIGSGVOEI	17
		1:11:11:11	
DB	107	RVIGISGLOEL	117

Search completed: October 1, 2003, 10:37:53  
Job time : 25.1176 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 ; Search time 54.4118 Seconds  
(without alignments)  
72.928 Million cell updates/sec

Title: US-09-171-432a-47  
Perfect score: 129  
Sequence: 1 ORLKYAQLSELSNEVLPPPKMKGLF 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 1107863 seqs, 158726573 residues

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: A.Geneseq.19Jun03: \*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: \*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: \*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: \*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: \*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: \*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: \*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: \*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: \*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: \*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: \*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: \*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: \*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: \*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: \*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: \*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: \*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: \*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: \*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: \*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: \*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: \*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: \*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: \*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	25	18 AAW42930	Immunogenic Hepat
2	129	100.0	26	22 AAB69447	Synthetic HAV P2A
3	129	100.0	366	6 AAF50230	Sequence of hepat
4	129	100.0	854	6 AAF50287	Sequence encoded b
5	129	100.0	993	6 AAF50116	Sequence of Hepat
6	129	100.0	993	6 AAF50231	Sequence encoded b
7	129	100.0	1077	20 AAW95559	A partial hepatitis
8	129	100.0	1091	14 AAR32426	Translated from 5'
9	129	100.0	2227	11 AAR05697	Attenuated hepatitis

10	129	100.0	2227	18 AAW34074	Hepatitis A virus
11	129	100.0	2227	21 AAB18607	Amino acid sequenc
12	129	100.0	2227	21 AAB18609	Amino acid sequenc
13	129	100.0	2227	23 AAB31727	Wild-type Hepatit
14	129	100.0	2227	23 AAB31729	Attenuated Hepatit
15	129	100.0	2227	23 AAB19899	Hepatitis A virus
16	129	100.0	2227	24 ABO08639	Hepatitis A virus
17	129	100.0	2227	24 ABO08641	Wild type human he
18	124	96.1	2227	7 AAB60066	Attenuated hepati
19	124	96.1	2227	23 AAB18608	Sequence of viral
20	124	96.1	2227	23 AAB31728	Amino acid sequenc
21	124	96.1	2227	24 ABO08640	Hepatitis A virus
22	117	90.7	839	12 AAB15629	Attenuated (pass3
23	103	79.8	20	18 AAW42924	Capsid region of c
24	103	79.8	21	22 AAB69441	Immunogenic Hepat
25	65	50.4	20	18 AAW42925	Synthetic HAV P2A
26	65	50.4	21	22 AAB69442	Immunogenic Hepat
27	51	39.5	150	22 AAB68507	Synthetic HAV P2A
28	49	38.0	440	22 AAB63362	Drosophila melanog
29	49	38.0	444	18 AAW20624	Drosophila melanog
30	48	37.2	187	21 AAB32794	H. pylori cytoplas
31	48	37.2	221	22 AAB68504	Eucalyptus grandis
32	48	37.2	1176	22 AAB59692	Drosophila melanog
33	47.5	36.8	86	21 AAG03559	Drosophila melanog
34	47.5	36.8	530	21 AAB42529	Human secreted pro
35	47.5	36.8	530	23 AAO26501	Human ORF2293
36	47.5	36.8	530	23 AAB68940	Human glycoprotein
37	46.5	36.0	23	12 AAR15517	Human polyprotein
38	46.5	36.0	137	14 AAR53388	BMP-8 peptide. Bo
39	46.5	36.0	139	14 AAR27288	Murine osteogenic
40	46.5	36.0	139	14 AAR33401	Mature murine oste
41	46.5	36.0	139	14 AAR33401	Murine mature OP-2
42	46.5	36.0	139	15 AAR46735	Morphogen mop-2.
43	46.5	36.0	139	15 AAR60970	Mouse osteogenic p
44	46.5	36.0	139	17 AAW00224	Murine OP-2
45	46.5	36.0	139	18 AAW40183	Mouse mature osteo
					Mouse mature osteo
					Mouse hOP2 mature

## ALIGNMENTS

RESULT 1					
ID	AAW42930	standard; peptide; 25 AA.			
XX					
AC	AAW42930:				
XX					
DT	28-APR-1998	(first entry)			
XX					
DE	Immunogenic Hepatitis A virus peptide YK-1665.				
XX					
KW	Immunogenic peptide; Immunogenic epitope; P2A protein;				
XX	Immune response; antibody.				
OS	Synthetic.				
OS	Hepatitis A virus.				
PN	WO9740147-A1.				
XX					
PD	30-OCT-1997.				
XX					
PF	18-APR-1997;	97WO-US06891.			
XX					
PR	19-APR-1996;	96US-0015644.			
XX					
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.				
XX					
PI	Fields HA, Khudiyakov YE;				
DR					
XX	WPI; 1997-535831/49.				
XX					
PT	Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an				
PT	Immune response to HAV in a mammal or to detect the presence of				

PT antibodies against HAV in a mammal  
XX  
PS Claim 18; Page 112; 140pp; English.  
XX  
CC Peptides AAW42922-30 are immunogenic peptides corresponding to  
CC immunogenic epitopes of the Hepatitis A virus (HAV). The peptides are  
CC substantially similar to a portion of the amino acid sequence of the P2A  
CC protein of HAV corresponding to amino acids 792-980. Compositions  
CC containing the peptides can be used to induce an immune response to HAV  
CC in a mammal. The peptides can also be used to detect the presence of  
CC antibodies against HAV in mammalian serum. The peptides can also be used  
CC to make an antibody against HAV by administering the peptide to a  
CC mammal.  
XX  
SQ Sequence 25 AA;  
XX  
Query Match 100.0%; Score 129; DB 18; Length 25;  
Best Local Similarity 100.0%; Pred. No. 8.3e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 QRLKYAOEELSNEVLPPEPRKKMGLE 25  
1 QRLKYAOEELSNEVLPPEPRKKMGLE 25  
RESULT 2  
AAB69447  
ID AAB69447 standard; Peptide: 26 AA.  
XX  
AC AAB69447;  
XX  
DT 20-APR-2001 (first entry)  
XX  
DE Synthetic HAV P2A peptide, SEQ ID NO: 47.  
XX  
KW Hepatitis A virus; HAV; immunogen; immunostimulant; virucide; vaccine;  
KW antigen; major structural capsid polypeptide; HAV antibody detection.  
XX  
OS Hepatitis A virus.  
OS Synthetic.  
XX  
PN WO200105824-A2.  
XX  
PD 25-JAN-2001.  
XX  
PF 14-JUL-2000; 2000WO-US19267.  
XX  
PR 15-JUL-1999; 99US-0144412.  
XX  
(USSH ) US DEPT HEALTH & HUMAN SERVICES.  
PI Fields HA, Khudyakov YE;  
XX  
DR WPI: 2001-112681/12.  
XX  
PT Synthetic peptides used as antigen sources for enzyme immunoassays  
PT detecting anti-hepatitis A virus and as vaccines -  
XX  
PS Claim 13; Page 98; 130pp; English.  
XX  
CC The present sequence is one of a number of synthetic peptides which are  
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides  
CC comprise antigenic epitopes of the major structural capsid polypeptides  
CC or non-structural polypeptides of HAV with one or more glutamine  
CC molecules at the carboxy end of the peptide. The peptides are used to  
CC detect the presence of antibodies against HAV in mammalian serum, to  
CC detect the presence of HAV in a human or animal through the binding of  
CC the peptide to an antibody, to detect acute phase infection by detecting  
CC IgM antibodies in mammalian serum and detecting convalescence in a  
CC mammal. The peptides are used to detect or quantify HAV antibodies in  
CC samples in clinical or research-based assays using immunoblotting,  
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,  
CC tracking of radioactive or bioluminescent markers, chromatography or

CC electrophoresis. The peptides are used to induce an immune response to  
CC HAV when administered to a human or animal. Glutamine at the carboxy  
CC end of the peptides enhances the IgM antibody reactivity.  
XX  
SQ Sequence 26 AA;  
XX  
Query Match 100.0%; Score 129; DB 22; Length 26;  
Best Local Similarity 100.0%; Pred. No. 8.7e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 QRLKYAOEELSNEVLPPEPRKKMGLE 25  
1 QRLKYAOEELSNEVLPPEPRKKMGLE 25  
RESULT 3  
AAP50230  
ID AAP50230 standard; Protein: 366 AA.  
XX  
AC AAP50230;  
XX  
DT 28-NOV-1991 (first entry)  
XX  
DE Sequence of hepatitis A virus (HAV) surface protein (VP-1).  
XX  
KW Hepatitis A virus vaccine; immunisation; monoclonal antibody;  
KW diagnostic assay.  
XX  
OS Hepatitis A virus.  
XX  
PN EPI38704-A.  
XX  
PD 24-APR-1985.  
XX  
PF 09-OCT-1984; 84EP-0402025.  
XX  
PR 02-MAR-1984; 84US-0585942.  
XX  
PR 14-OCT-1983; 83US-0541836.  
XX  
PA (MERI ) MERCK & CO INC.  
XX  
PI Hughes JV, Scolnick EM, Tomassini JE;  
XX  
DR WPI: 1985-100818/17.  
XX  
DR N-PSDB; AAN50274.  
XX  
PT New hepatitis A virus surface protein - useful for binding to  
PT neutralising antibodies to the virus  
XX  
PS Claim 21; Page 46-48; 49pp; English.  
XX  
CC VPI is isolated by solubilisation of the intact virus in an aq.  
CC anionic surfactant and a reducing agent. The viral proteins are sepd.  
CC and the protein of molecular wt. 33000 daltons is sepd.  
XX  
SQ Sequence 366 AA;  
XX  
Query Match 100.0%; Score 129; DB 6; Length 366;  
Best Local Similarity 100.0%; Pred. No. 1.5e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
DB 1 QRLKYAOEELSNEVLPPEPRKKMGLE 25  
1 QRLKYAOEELSNEVLPPEPRKKMGLE 343  
RESULT 4  
AAP50287  
ID AAP50287 standard; Protein: 854 AA.  
XX  
AC AAP50287;  
XX  
DT 25-MAR-2003 (updated)

DT 30-NOV-1991 (first entry)  
XX  
DE Sequence encoded by hepatitis A virus (HAV) cDNA from near the  
DE genome 5' terminus to the end of the area corresponding to the  
DE capsid protein region of poliovirus RNA.  
XX  
XX Hepatitis A virus assay; antigen; antibody.  
XX  
XX Hepatitis A virus.  
XX  
XX W05501517-A.  
XX  
XX 11-APR-1985.  
XX  
XX 27-SEP-1984; 84MO-US01552.  
XX  
XX 30-SEP-1983; 83US-0537911.  
XX  
XX (MAST ) MASSACHUSETTS INST TECHNOLOGY.  
XX  
XX Ticehurst JR, Baltimore D, Feinstein SM, Purcell RH;  
XX Racanello VR;  
XX WPI: 1985-098846/16.  
XX N-PSDB: AAN50330.  
XX  
XX New hepatitis A virus cDNA - useful in assays for the virus and  
XX for prodn. of the viral antigen and antibodies to it  
XX  
XX Example: Fig 7: 60pp: English.  
XX  
XX The inventors claim HAV cDNA and a method for producing it, whereby  
XX large amts. can be obt'd. economically. The cDNA is useful in the  
XX assay for detection of HAV quickly and easily and with high  
XX sensitivity and specificity. The HAV cDNA is also used in the prodn.  
XX of HAV antigen or antibodies to it. The antibodies may be monoclonal.  
XX (Updated on 25-MAR-2003 to correct PA field.)  
XX  
XX Sequence 854 AA:  
SQ  
Query Match 100.0%; Score 129; DB 6; Length 854;  
Best Local Similarity 100.0%; Pred. No. 3.8e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ORLKYAOEELSNEVLPPPRKMGFLF 25  
DB 810 ORLKYAOEELSNEVLPPPRKMGFLF 834  
RESULT 5  
AAP50116  
ID AAP50116 standard; Protein; 993 AA.  
XX  
XX AAP50116;  
XX  
XX 25-MAR-2003 (updated)  
XX 30-SEP-1991 (first entry)  
XX  
XX Sequence of Hepatitis A virus (HAV) immunogenic peptides  
XX VP-1, VP-2, VP-3 and VP-4.  
XX  
XX Antigenic protein; immunogen; vaccine.  
XX  
XX Hepatitis A virus (strain CR326).  
XX  
XX EPI54587-A.  
XX  
XX 11-SEP-1985.  
XX  
XX 27-FEB-1985; 85EP-0400369.  
XX  
XX 02-MAR-1984; 84US-0585818.  
XX

PA (MERI ) MERCK & CO INC.  
XX  
XX Linemeyer DL, Menke JG, Rueben RG, Miltra SM;  
XX WPI: 1985-224964/37.  
XX N-PSDB: AAN50139.  
XX  
XX New nucleotide sequences coding for hepatitis A virus antigens -  
XX useful for eliciting normal immune response and in vaccines for  
XX protecting against the virus  
XX  
XX Example: Page 11-17; 32pp; English.  
XX  
XX Within the sequence in AAN50139 is encoded the information necessary  
XX to make the antigenic proteins of HAV. The sequences encoding for  
XX the structural proteins begin at base 403. The key sub-unit  
XX sequences within VP-1, designated Sequences I,II,III,IV, and V,  
XX start, respectively at 1882, 1963, 1999, 2146, 2347. Other  
XX nucleotide sequences which are valuable as encoding antigenic  
XX proteins are the sequences from base 1749 to base 2722; from base  
XX 1487 to base 2980 and from base 1644 to base 2722. The sequence from  
XX base 1749 to base 2722 is esp. valuable as a vector for producing the  
XX antigen protein. Sequences II-V are claimed. X in AAP50116 denotes the  
XX translation of a stop codon.  
XX  
XX (Updated on 25-MAR-2003 to correct PI field.)  
XX  
XX Sequence 993 AA:  
SQ  
Query Match 100.0%; Score 129; DB 6; Length 993;  
Best Local Similarity 100.0%; Pred. No. 4.4e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ORLKYAOEELSNEVLPPPRKMGFLF 25  
DB 946 ORLKYAOEELSNEVLPPPRKMGFLF 970  
RESULT 6  
AAP50231  
ID AAP50231 standard; Protein; 993 AA.  
XX  
XX AAP50231;  
XX  
XX 28-NOV-1991 (first entry)  
XX  
XX Sequence encoded by partial sequence of hepatitis A virus (HAV),  
XX including surface protein (VP-1).  
XX  
XX Hepatitis A virus vaccine; immunisation; monoclonal antibody;  
XX diagnostic assay.  
XX  
XX Hepatitis A virus.  
XX  
XX Key Location/Qualifiers  
XX FT Protein 628..993  
XX FT /note="claimed; X denotes translated stop codons  
XX FT and unspecified triplets"  
XX  
XX EPI38704-A.  
XX  
XX 24-APR-1985.  
XX  
XX 09-OCT-1984; 84EP-0402025.  
XX  
XX 02-MAR-1984; 84US-0585942.  
XX 14-OCT-1983; 83US-0541836.  
XX  
XX (MERI ) MERCK & CO INC.  
XX  
XX Hughes JV, Scolnick EM, Tomassini JB;  
XX WPI: 1985-100818/17.  
XX N-PSDB: AAN50274.  
XX

```

XX  New hepatitis A virus surface protein - useful for binding to
PT  neutralising antibodies to the virus
XX
XX  Disclosure; Page 17-23; 49pp; English.
XX
CC  VPI is isolated by solubilisation of the intact virus in an aq.
CC  anionic surfactant and a reducing agent. The viral proteins are sepd.
CC  and the protein of molecular wt. 33000 daltons is sepd.
XX
SQ  Sequence 993 AA:

Query Match      100.0%; Score 129; DB 6; Length 993;
Best Local Similarity 100.0%; Pred. No. 4.4e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  1 QRLKYAOEELSNEVLPPEPRKMGLE 25
    |||||
DB  946 QRLKYAOEELSNEVLPPEPRKMGLE 970

ID  AAW95559 standard; Protein; 1077 AA.
XX
AC  AAW95559;
XX
DT  28-APR-1999 (first entry)
XX
DE  A partial hepatitis A virus (HAV) protein.
XX
KM  Hepatitis A virus protein; HAV; P2 region;
KW  cell-culture-adapted HAV strain; infection; accelerated growth.
XX
OS  Hepatitis A virus.
XX
PN  US5849562-A.
XX
PD  15-DEC-1998.
XX
PF  06-JUN-1995; 95US-0468926.
XX
PR  06-NOV-1991; 91US-0788262.
PR  30-SEP-1983; 83US-0537911.
PR  27-SEP-1984; 84US-0654942.
PR  06-OCT-1988; 88US-0256135.
PR  06-JUN-1995; 95US-0468926.
XX
XX  (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX  Emerson SU, Purcell RH;
XX
XX  WPI; 1999-094412/08.
XX
XX  N-PSDB; AAX01006.
XX
PT  Chimeric hepatitis A virus strains - with P2 region from
PT  cell-culture-adapted strain in wild-type genome
XX
PS  Disclosure; Fig 7A-L; 36pp; English.
XX
XX  The present sequence represents a partial hepatitis A virus (HAV)
XX  protein. The specification describes a DNA construct consisting
XX  of a wild-type HAV genome in which the P2 region is replaced by the
XX  P2 region from a cell-culture-adapted HAV strain. The construct is
XX  used to demonstrate that mutations in the P2 region of a
XX  cell-culture-adapted HAV strain are sufficient for establishment of
XX  infection and accelerated growth in cell culture.
XX
SQ  Sequence 1077 AA:

Query Match      100.0%; Score 129; DB 20; Length 1077;
Best Local Similarity 100.0%; Pred. No. 4.8e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY  1 QRLKYAOEELSNEVLPPEPRKMGLE 25
    |||||
DB  1033 QRLKYAOEELSNEVLPPEPRKMGLE 1057

RESULT 8
AAR32426
ID  AAR32426 standard; Protein; 1091 AA.
XX
AC  AAR32426;
XX
XX  25-MAR-2003 (updated)
DT  17-DEC-2001 (updated)
DT  10-JUN-1993 (first entry)
XX
DE  Translated from 5' region of Hepatitis A virus genomic clone.
XX
XX  HAV HM-175; chronic liver disease; picornavirus.
XX
XX  Hepatitis A virus.
XX
XX  OS
XX  FH
XX  Key      Location/Qualifiers
XX  Region   238..1091
XX           /label= ORF
XX           /note= "second putative initiation codon at
XX           position 240"
XX  Region   1..711
XX           /note= "X's correspond to nonsense codons,
XX           1.e. this region is not an ORF"
XX
XX  USN7788262-N.
XX
XX  15-DEC-1992.
XX
XX  30-SEP-1983; 83US-0536911.
XX
XX  27-SEP-1984; 84US-0654942.
XX  06-OCT-1988; 88US-0256135.
XX  30-SEP-1983; 83US-0536911.
XX  06-NOV-1991; 91US-0788262.
XX
XX  (USSH ) US DEPT HEALTH & HUMAN SERVICE.
XX
XX  Ricehurst JR, Baltimore D, Feinstein SM, Purcell RH, Racanietello VR;
XX  Baroudy BM, Emerson SU;
XX
XX  WPI: 1993-067429/08.
XX
XX  N-PSDB; AAO36934.
XX
XX  Hepatitis A virus cDNA prodn. - for diagnostic use and for prodn.
XX  of antigen and antibodies
XX
XX  Disclosure; Fig 7; 65pp; English.
XX
XX  HAV virion RNA was extracted from the livers of marmosets which had
XX  been inoculated with HAV (the HAV had previously been passaged twice
XX  in marmosets). The RNA was used to prepare ds cDNA clones by
XX  standard methods. Clones contg. inserts which hybridised to RNA from
XX  HAV-infected African Green Monkey kidney cells were selected for
XX  further analysis. A 7.4kb restriction map (about 9% of the HAV
XX  genome) was constructed from 5 overlapping inserts. The sequence of
XX  the first 3.3kb (approx.) from the 5'-terminus was determined. An
XX  amino acid sequence was deduced from the entire clone and an open
XX  reading frame was identified starting at position 238. A comparison
XX  of the predicted HAV amino acid sequences with the known capsid
XX  protein sequences of other picornaviruses (poliovirus, foot and
XX  mouth disease virus and encephalomyelitis virus) revealed areas of
XX  local homology.
XX
XX  (Note: Revised entry submitted to correct the patent number format of
XX  US Government-owned NRTS applications to prevent clashes with ongoing US
XX  granted patent numbers. For further information please visit the Derwent
XX  web site at www.derwent.com/dwpi/updates/nrtis-us.html.)

```

CC (Updated on 25-MAR-2003 to correct PI field.)  
XX  
SQ Sequence 1091 AA;  
Query Match 100.0%; Score 129; DB 14; Length 1091;  
Best Local Similarity 100.0%; Pred. No. 4.9e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ORLKYAOEELSNEVLPPEPRKMGLE 25  
DB 1047 ORLKYAOEELSNEVLPPEPRKMGLE 1071  
RESULT 9  
AAR05697  
ID AAR05697 standard; protein; 2227 AA.  
XX  
AC AAR05697;  
XX  
XX 25-MAR-2003 (updated)  
XX 15-AUG-1990 (first entry)  
DE Attenuated hepatitis A virus.  
XX  
XX Hepatitis A virus; vaccine; attenuated.  
XX  
OS Hepatitis A virus, strain HM-175.  
XX  
XX  
FH Key Location/Qualifiers  
FT 1..23  
FT /label-VP4 = 1A  
FT 24..245  
FT /label-VP2 = 1B  
FT 246..491  
FT /label-VP3 = 1C  
FT 492..791  
FT /label-VP1 = 1D  
FT 792..980  
FT /label-2A  
FT 981..1087  
FT /label-2B  
FT 1088..1422  
FT /label-2C  
FT 1423..1496  
FT /label-3A  
FT 1497..1519  
FT /label-3B = VPg  
FT 1520..1738  
FT /label-3C  
FT 1739..2227  
FT /label-3D  
XX  
XX  
XX US4894228-A.  
XX  
XX  
XX 16-JAN-1990.  
XX  
XX  
XX 12-JUL-1988; 88US-0217824.  
XX  
XX 12-JUL-1988; 88US-0217824.  
XX 12-JUL-1988; 88US-0652967.  
XX  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
XX  
XX Purcell RH, Ticehurst JR, Cohen JL, Emerson SU, Feinstein SM;  
XX Daemer RJ, Gust ID;  
XX  
XX WPI: 1990-075557/10.  
XX N-PSDB: AAQ03512.  
XX  
XX Vaccine against hepatitis A virus infection comprises novel  
XX attenuated hepatitis A virus strain.  
XX  
XX Claim 1; Fig 1; 18pp; English.  
XX  
XX

XX  
XX The attenuated HAV is useful for inducing protective immunity against  
XX HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by  
XX several nucleotide changes distributed throughout the genome, is  
XX attenuated for chimpanzees, elicits serum neutralising antibodies, and is  
XX suitable for use as an HAV vaccine. It is noted that not all the changes  
XX are necessary for attenuation and use as a vaccine.  
XX (Updated on 25-MAR-2003 to correct PI field.)  
XX  
CC (Updated on 25-MAR-2003 to correct PI field.)  
CC  
SQ Sequence 2227 AA;  
Query Match 100.0%; Score 129; DB 11; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.1e-10;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ORLKYAOEELSNEVLPPEPRKMGLE 25  
DB 810 ORLKYAOEELSNEVLPPEPRKMGLE 834  
RESULT 10  
AAW34074  
ID AAW34074 standard; Protein; 2227 AA.  
XX  
XX  
AC AAW34074;  
XX  
XX  
DT 27-APR-1998 (first entry)  
XX  
XX  
DE Hepatitis A virus HM-175 protein sequence.  
XX  
XX HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;  
XX infection; vaccine.  
XX  
OS Hepatitis A virus HM-175.  
XX  
XX  
FH Key Location/Qualifiers  
FT 1..23  
FT /label-VP4  
FT 24..245  
FT /label-VP2  
FT 246..491  
FT /label-VP3  
FT 492..791  
FT /label-VP1  
FT 792..980  
FT /label-2A  
FT 981..1087  
FT /label-2B  
FT 1088..1422  
FT /label-2C  
FT 1423..1496  
FT /label-3A  
FT 1497..1519  
FT /label-3B  
FT 1520..1738  
FT /label-3C  
FT 1739..2227  
FT /label-3D  
XX  
XX  
XX WO9740166-A2.  
XX  
XX  
XX 30-OCT-1997.  
XX  
XX  
XX 18-APR-1997; 97WO-US06506.  
XX  
XX 19-APR-1996; 96US-0015642.  
XX  
XX  
XX (USSH ) US SEC DEPT HEALTH.  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Emerson SU, Purcell RH, Raychaudhuri G;  
XX  
XX

DR WPI: 1997-535850/49.  
XX N-PSDB: AAT93023.  
XX  
PT Human attenuated HAV genome containing simian HAV 2C gene - useful  
PT as vaccine against HAV infection  
XX  
PS Disclosure: Fig 13A-D; 66pp; English.  
XX  
CC This protein sequence is encoded by the human hepatitis A virus  
CC (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain  
CC HAV/7 is obtained by passage of HM-175 in African Green Monkey  
CC kidney cells. A claimed DNA construct (I) comprises a genome of  
CC HAV, where the genome is a human attenuated HAV genome in which a  
CC region of the 2C gene has been replaced by a corresponding region  
CC from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The  
CC region of the 2C gene from AGM-27 contained in the construct  
CC preferably encodes amino acids 120-328 of the 2C protein, amino  
CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA  
CC transcript of (1); (2) a cell transfected with (1) or the RNA  
CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the  
CC HAV of (3); and (5) a host cell containing the HAV of (3). (1) or  
CC its RNA transcript, can be used as a vaccine for preventing HAV in  
CC a mammal. (1) or the RNA transcript can also be used to stimulate  
CC the production of protective antibodies in the mammal.  
CC  
XX  
SQ Sequence 2227 AA:  
Query Match 100.0%; Score 129; DB 18; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.1e-10;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ORLKYAOEELSNEVLPPrKKMGLE 25  
DB 810 ORLKYAOEELSNEVLPPrKKMGLE 834  
RESULT 11  
AAB18607  
ID AAB18607 standard; Protein; 2227 AA.  
XX  
AC AAB18607;  
XX  
DT 15-JAN-2001 (first entry)  
XX  
DE Amino acid sequence of wild type Hepatitis A virus strain HM-175.  
XX  
KW HAV, strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.  
XX  
XX Hepatitis A virus.  
XX  
XX US6113912-A.  
XX  
PD 05-SEP-2000.  
XX  
PF 07-JUN-1995; 95US-0475886.  
XX  
PR 18-SEP-1992; 92US-0947338.  
PR 17-SEP-1993; 93WO-US08610.  
PR 10-MAR-1995; 95US-0397232.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;  
XX  
XX WPI: 2000-586464/55.  
XX  
XX N-PSDB: AAA75476.  
XX  
PT Novel live hepatitis A virus adapted to growth in human fibroblast cell  
PT line useful as vaccine for protecting humans against hepatitis A virus  
PT infection, has modified genome compared to wild type -  
XX  
PS Disclosure: Fig 6A-K; 72pp; English.  
XX

CC The present sequence is derived from a wild type hepatitis A virus  
CC (HAV) strain HM-174. The sequence is modified to produce HAV which  
CC are adapted to growth in the human fibroblast-like cell line MRC-5.  
CC The HAV is able to propagate in MRC-5 cells and retain appropriate  
CC attenuation. It is useful as a live vaccine for prophylaxis of  
CC hepatitis A in humans and other primates.  
XX  
XX  
SQ Sequence 2227 AA:  
Query Match 100.0%; Score 129; DB 21; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.1e-10;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ORLKYAOEELSNEVLPPrKKMGLE 25  
DB 810 ORLKYAOEELSNEVLPPrKKMGLE 834  
RESULT 12  
AAB18609  
ID AAB18609 standard; Protein; 2227 AA.  
XX  
AC AAB18609;  
XX  
DT 15-JAN-2001 (first entry)  
XX  
DE Amino acid sequence of live attenuated Hepatitis A virus 4380.  
XX  
KW HAV, strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;  
XX HAV 4380.  
XX  
XX Hepatitis A virus.  
XX  
XX US6113912-A.  
XX  
PD 05-SEP-2000.  
XX  
PF 07-JUN-1995; 95US-0475886.  
XX  
PR 18-SEP-1992; 92US-0947338.  
PR 17-SEP-1993; 93WO-US08610.  
PR 10-MAR-1995; 95US-0397232.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;  
XX  
XX WPI: 2000-586464/55.  
XX  
XX N-PSDB: AAA75478.  
XX  
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell  
XX line useful as vaccine for protecting humans against hepatitis A virus  
XX infection, has modified genome compared to wild type -  
XX  
XX  
PS Disclosure: Columns 93-104; 72pp; English.  
XX  
XX The present sequence is derived from a live attenuated hepatitis A  
XX virus (HAV) of the invention, designated HAV 4380. The sequence is  
XX produced by modifying wild type HAV strain HM-174. The HAV of the  
XX invention are adapted to growth in the human fibroblast-like cell  
XX line MRC-5. The HAV is able to propagate in MRC-5 cells and retain  
XX appropriate attenuation. It is useful as a live vaccine for prophylaxis  
XX of hepatitis A in humans and other primates.  
XX  
XX  
SQ Sequence 2227 AA:  
Query Match 100.0%; Score 129; DB 21; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.1e-10;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ORLKYAOEELSNEVLPPrKKMGLE 25  
DB 810 ORLKYAOEELSNEVLPPrKKMGLE 834



RESULT 13  
 ABG31727  
 ID ABG31727 standard; Protein: 2227 AA.  
 XX  
 AC ABG31727;  
 XX  
 DT 29-NOV-2002 (first entry)  
 XX  
 DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.  
 XX  
 KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.  
 XX  
 OS Hepatitis A virus strain HM-175.  
 XX  
 PN US6423318-B1.  
 PD 23-JUL-2002.  
 31-AUG-2000; 2000US-0653499.  
 07-JUN-1995; 95US-0475886.  
 PR 17-SEP-1993; 93US-0397232.  
 PR 17-SEP-1993; 93WO-US08610.  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Funckhouser AM, Emerson SU, Purcell RH, D'Hondt E;  
 DR WPI: 2002-680946/73.  
 DR N-PSDB; ABS52787.  
 XX  
 PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth  
 in MRC-5 cells, useful for preparing a vaccine against HAV infection -  
 XX  
 PS Disclosure; Fig 6; 71pp; English.  
 XX  
 CC The invention relates to a polynucleotide which encodes a hepatitis A  
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell  
 CC line). The polynucleotide is useful for preparing a vaccine against  
 CC hepatitis A virus infection. This sequence represents a hepatitis A virus  
 CC strain HM-175 polypeptide.  
 CC  
 SQ Sequence 2227 AA:  
 Query Match 100.0%; Score 129; DB 23; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 1,1e-10;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 ORLKYAQEELSNEVLP PPRKMGLE 25  
 DB 810 ORLKYAQEELSNEVLP PPRKMGLE 834  
 RESULT 14  
 ABG31729  
 ID ABG31729 standard; Protein: 2227 AA.  
 XX  
 AC ABG31729;  
 XX  
 DT 29-NOV-2002 (first entry)  
 XX  
 DE Attenuated Hepatitis A virus (HAV) 4380 polypeptide.  
 XX  
 KW Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;  
 KW HAV 4380.  
 XX  
 OS Hepatitis A virus strain HM-175.  
 XX  
 PN US6423318-B1.  
 PD

PD 23-JUL-2002.  
 XX  
 PF 31-AUG-2000; 2000US-0653499.  
 XX  
 PR 07-JUN-1995; 95US-0475886.  
 PR 17-SEP-1993; 93US-0397232.  
 PR 17-SEP-1993; 93WO-US08610.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.  
 XX  
 PI Funckhouser AM, Emerson SU, Purcell RH, D'Hondt E;  
 DR WPI: 2002-680946/73.  
 DR N-PSDB; ABS52789.  
 XX  
 PT New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth  
 in MRC-5 cells, useful for preparing a vaccine against HAV infection -  
 XX  
 PS Disclosure; Column 93-104; 71pp; English.  
 XX  
 CC The invention relates to a polynucleotide which encodes a hepatitis A  
 CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell  
 CC line). The polynucleotide is useful for preparing a vaccine against  
 CC hepatitis A virus infection. This sequence represents an attenuated  
 CC hepatitis A virus 4830 polypeptide.  
 CC  
 SQ Sequence 2227 AA:  
 Query Match 100.0%; Score 129; DB 23; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 1,1e-10;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 ORLKYAQEELSNEVLP PPRKMGLE 25  
 DB 810 ORLKYAQEELSNEVLP PPRKMGLE 834  
 RESULT 15  
 AAEL1989  
 ID AAEL1989 standard; Protein: 2227 AA.  
 XX  
 AC AAEL1989;  
 XX  
 DT 18-JUN-2002 (first entry)  
 XX  
 DE Hepatitis A virus (HAV) protein.  
 XX  
 KW Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;  
 KW cytostatic; immunostimulant; vaccine; ribavirin; immune response; cancer.  
 XX  
 OS Hepatitis A virus.  
 XX  
 PN WO200213855-A2.  
 PD 21-FEB-2002.  
 XX  
 PF 15-AUG-2001; 2001WO-1B01808.  
 XX  
 PR 17-AUG-2000; 2000US-225767P.  
 PR 29-AUG-2000; 2000US-229175P.  
 PR 03-NOV-2000; 2000US-0705547.  
 XX  
 PA (TRIP-) TRIPEP AB.  
 XX  
 PI Sallberg M, Hultgren C;  
 DR WPI: 2002-241837/29.  
 DR N-PSDB; AAD31766.  
 XX  
 PT Vaccine compositions for treating and preventing disease, preferably  
 PT hepatitis C virus infection, comprises ribavirin and antigen that has  
 PT epitope present in hepatitis C virus

XX Claim 11; Page 82-87; 120pp; English.  
 PS  
 CC The invention relates to a composition comprising ribavirin and an  
 CC antigen preferably non structural 3 protein (NS3)/4A fragment of  
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV  
 CC sequence. The composition is useful for enhancing an immune response to  
 CC a hepatitis C antigen in humans, domestic, sport or pet species and as  
 CC vaccines for treating and preventing HCV infections. The composition is  
 CC also useful for treating viral, bacterial, fungal diseases and cancer.  
 CC The present sequence is hepatitis A virus (HAV) protein.  
 XX  
 SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 23; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-10;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 QRLKYAOELSNLYLPPPRKMKGLF 25  
 ||||||||||||||||||||||||  
 810 QRLKYAOELSNLYLPPPRKMKGLF 834

Search completed: October 1, 2003, 09:56:46  
 Job time : 55.4118 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 : Search time 16.0294 Seconds  
(without alignments)  
149.988 Million cell updates/sec

Title: US-09-171-432a-47

Perfect score: 129

Sequence: 1 QRLRYAQBELSNVLPPEPRKMKGLF 25

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

al number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: pIR1:\*  
2: pIR2:\*  
3: pIR3:\*  
4: pIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	341	2 S04137	genome polypeptide
2	129	100.0	852	1 GNNYHA	genome polypeptide
3	129	100.0	1358	2 A03905	genome polypeptide
4	129	100.0	2227	1 GNNYHM	genome polypeptide
5	129	100.0	2227	1 GNNYHM	genome polypeptide
6	124	96.1	2227	1 GNNYHM	genome polypeptide
7	123	95.3	2227	1 GNNYHM	genome polypeptide
8	119	92.2	2230	1 GNNYSA	genome polypeptide
9	117	90.7	839	1 GNNYSA	genome polypeptide
10	56.5	43.8	346	2 S74448	genome polypeptide
11	52	40.3	859	2 T43701	regulatory protein
12	52	40.3	1193	2 E88445	DNA-directed RNA P
13	51	39.5	1119	2 T50995	protein C26E6.4 (1
14	51	39.5	6642	2 T29757	related to cytoske
15	50.5	39.1	443	2 E82046	protein UNC-89 - C
16	49.5	38.4	159	2 C72210	proteinase HsIVU
17	49.5	38.4	554	2 A56730	conserved hypochet
18	49	38.0	55	2 P00433	carl protein - Pod
19	49	38.0	55	2 P00434	genome polypeptide
20	49	38.0	56	2 P00428	genome polypeptide
21	49	38.0	56	2 P00427	genome polypeptide
22	49	38.0	56	2 P00429	genome polypeptide
23	49	38.0	56	2 P00432	genome polypeptide
24	49	38.0	56	2 P00430	genome polypeptide
25	49	38.0	442	2 A71969	probable histidine
26	49	38.0	465	2 H86482	protein F5U5.11 (1
27	49	38.0	1174	2 S28976	DNA-directed RNA P
28	48.5	37.6	2278	1 S56274	FAB1 protein - yea
29	48	37.2	523	2 D8538	hypothetical prote

30	48	37.2	523	2 H90687	hypothetical prote
31	48	37.2	1176	2 A27826	DNA-directed RNA P
32	48	37.2	1191	2 S65068	DNA-directed RNA P
33	48	37.2	1210	2 S35548	DNA-directed RNA P
34	47.5	36.8	285	2 T15133	hypothetical prote
35	47.5	36.8	1626	2 T09271	probable tail comp
36	47	36.4	223	2 P00514	hemagglutinin HA2
37	47	36.4	254	2 F75575	hypothetical prote
38	47	36.4	578	2 S03299	hemagglutinin prec
39	47	36.4	592	1 LIRY	actin-binding prote
40	46.5	36.0	361	2 G82530	hypothetical prote
41	46.5	36.0	486	2 T10089	clitrin - sweet ora
42	46.5	36.0	634	2 E86293	T24D18.1 protein -
43	46	35.7	163	2 PC4186	hemagglutinin 2 ch
44	46	35.7	262	2 A70428	hypothetical prote
45	46	35.7	338	2 I56893	transcription fact

## ALIGNMENTS

## RESULT 1

S04137 genome polypeptide - human hepatitis A virus (strain LCDC-1) (fragment)

C:Species: human hepatitis A virus

C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 21-Jul-2000

C:Accession: S04137

R:Anderson, A.P.; Lau, P.C.K.; Chaudhary, R.

Nucleic Acids Res. 17, 3594, 1989

A:Title: Nucleotide sequence of the VP1 gene from a Chinese strain of hepatitis A vir

A:Reference number: S04137; MID:89263805; PMID:2542903

A:Accession: S04137

A:Molecule type: mRNA

A:Residues: 1-341 <RAND>

A:Cross-references: EMBL:X14666; NID:g62301; PIDN:CAA32794.1; PID:g4377576

C:Genetics:

A:Gene: VP1

C:Superfamily: hepatitis A virus genome polypeptide

C:Keywords: coat protein; polypeptide

F:2-340/Product: coat protein ID (VP1) #status predicted <MAT>

Query Match	100.0%; Score 129; DB 2; Length 341;
Best Local Similarity	100.0%; Pred. No. 9.4e-12;
Matches 25; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
DB	314 QRLRYAQBELSNVLPPEPRKMKGLF 338

## RESULT 2

GNNYHA genome polypeptide - human hepatitis A virus (strain CR326) (fragment)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A

C:Species: human hepatitis A virus

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Jul-1999

C:Accession: A03904

R:Linemeyer, D.L.; Menke, J.G.; Martin-Gallardo, A.; Hughes, J.V.; Young, A.; Mitra,

J. Virol. 54, 247-255, 1985

A:Title: Molecular cloning and partial sequencing of hepatitis A viral cDNA.

A:Reference number: A03904; MID:85185648; PMID:2985793

A:Accession: A03904

A:Molecule type: genomic RNA

A:Residues: 1-852 <LIN>

A:Cross-references: EMBL:M10033; NID:g329592; PIDN:AAA45470.1; PID:g329593

C:Superfamily: hepatitis A virus genome polypeptide

C:Keywords: coat protein; core protein; polypeptide

F:1-245/Product: coat protein 1A #status predicted <CIA>

F:246-491/Product: coat protein 1B #status predicted <CIA>

F:492-836/Product: coat protein 1C #status predicted <CIA>

F:837-852/Product: core protein 2A (fragment) #status predicted <C2A>

Query Match 100.0%; Score 129; DB 1; Length 852;  
Best Local Similarity 100.0%; Pred. No. 2,7e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ORLKYAOEELSNEVLPPIPRKMKGLF 25  
|||||  
Db 810 ORLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 3  
genome polyprotein (version 2) - human hepatitis A virus (fragments)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; protein  
C; Species: human hepatitis A virus  
C:Date: 31-Mar-1991 #sequence, revision 31-Mar-1991 #text, change 15-Nov-1996  
Accession: A03905  
R:Baroudy, B.M.; Ticehurst, J.R.; Mele, T.A.; Maisel Jr., J.V.; Purcell, R.H.; Feinstor  
Proc. Natl. Acad. Sci. U.S.A. 82, 2143-2147, 1985  
A:Title: Sequence analysis of hepatitis A virus cDNA coding for capsid proteins and RNA  
A:Reference number: A03905; MUID:85166289; PMID:2984684  
Accession: A03905  
A:Molecule type: genomic RNA  
A:Residues: 1-1358 <BAR>

C:Superfamily: hepatitis A virus genome polyprotein  
C:Keywords: coat protein; core protein; polyprotein  
F:1-245/Product: coat protein 1A #status predicted <C1A>  
F:246-491/Product: coat protein 1B #status predicted <C1B>  
F:492-836/Product: coat protein 1C #status predicted <C1C>  
F:837-954/Product: core protein 2A (fragment) #status predicted <C2A>  
F:855-1358/Product: protein 3D (RNA polymerase) (fragment) #status predicted <C3D>

Query Match 100.0%; Score 129; DB 2; Length 1358;  
Best Local Similarity 100.0%; Pred. No. 4.6e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ORLKYAOEELSNEVLPPIPRKMKGLF 25  
|||||  
Db 810 ORLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 4  
GNNYHM  
genome polyprotein - human hepatitis A virus (strain HM-175, wild type)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core pro  
B; RNA-directed RNA polymerase (EC 2.7.7.48); protein 3D  
C:Species: human hepatitis A virus  
A:Note: host Homo sapiens (man)  
C:Date: 31-Mar-1988 #sequence, revision 31-Mar-1988 #text, change 16-Jul-1999  
Accession: A25981  
R:Chen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.  
Virol. 61, 50-59, 1987  
A:Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with d  
A:Reference number: A25981; MUID:87061253; PMID:3023706  
Accession: A25981  
A:Molecule type: genomic RNA  
A:Residues: 1-2227 <COH>  
A:Cross-references: EMBL:M14707; NID:9329582; PIDN:AAA45465.1; PID:9329583  
C:Superfamily: hepatitis A virus genome polyprotein  
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr  
F:1-23/Product: coat protein 1A #status predicted <VP4>  
F:24-245/Product: coat protein 1B #status predicted <VP2>  
F:246-491/Product: coat protein 1C #status predicted <VP3>  
F:492-791/Product: coat protein 1D #status predicted <VP1>  
F:792-980/Product: core protein 2A #status predicted <C2A>  
F:981-1087/Product: core protein 2B #status predicted <C2B>  
F:1088-1422/Product: core protein 2C #status predicted <C2C>  
F:1423-1496/Product: core protein 3A #status predicted <C3A>  
F:1497-1519/Product: protein 3B #status predicted <C3B>  
F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>  
F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 129; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 8e-11;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ORLKYAOEELSNEVLPPIPRKMKGLF 25  
|||||  
Db 810 ORLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 5  
genome polyprotein - human hepatitis A virus  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core  
NA polymerase (EC 2.7.7.48); protein 3D  
C:Species: human hepatitis A virus  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1987 #sequence, revision 30-Jun-1987 #text, change 16-Jul-1999  
Accession: A03903  
R:Netarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N  
Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985  
A:Title: Primary structure and gene organization of human hepatitis A virus.  
A:Reference number: A03903; MUID:85190549; PMID:2986137  
Accession: A03903  
A:Molecule type: genomic RNA  
A:Residues: 1-2227 <NAJ>  
A:Cross-references: GB:K02990; NID:9329596; PIDN:AAA45472.1; PID:9329597  
C:Superfamily: hepatitis A virus genome polyprotein  
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr  
F:1-245/Product: coat protein 1A #status predicted <C1A>  
F:246-491/Product: coat protein 1B #status predicted <C1B>  
F:492-836/Product: coat protein 1C #status predicted <C1C>  
F:837-980/Product: core protein 2A #status predicted <C2A>  
F:981-1076/Product: core protein 2B #status predicted <C2B>  
F:1077-1422/Product: core protein 2C #status predicted <C2C>  
F:1423-1484/Product: protein 3A #status predicted <C3A>  
F:1485-1507/Product: protein 3B #status predicted <C3B>  
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <C3C>  
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 129; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 8e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPIPRKMKGLF 25  
|||||  
Db 810 ORLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 6  
GNNYWK  
genome polyprotein - human hepatitis A virus (strain HM-175/7MK-5, attenuated HAV)  
N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; core protein 2A; core  
NA polymerase (EC 2.7.7.48); protein 3D  
C:Species: human hepatitis A virus  
A:Note: host Homo sapiens (man)  
C:Date: 30-Jun-1988 #sequence, revision 30-Jun-1988 #text, change 16-Jul-1999  
Accession: A94149; A25914; A94508  
R:Chen, J.I.; Rosenblum, B.; Ticehurst, J.R.; Daemer, R.J.; Feinstone, S.M.; Purcell  
Proc. Natl. Acad. Sci. U.S.A. 84, 2497-2501, 1987  
A:Title: Complete nucleotide sequence of an attenuated hepatitis A virus: comparison  
A:Reference number: A94149; MUID:87175701; PMID:3031686  
Accession: A94149  
A:Status: nucleic acid sequence not shown  
A:Molecule type: genomic RNA  
A:Residues: 1-2227 <COH>  
A:Cross-references: EMBL:M16632; NID:9329594; PIDN:AAA45471.1; PID:9329595  
A:Note: submitted to Genbank, August 1987  
C:Superfamily: hepatitis A virus genome polyprotein  
C:Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr  
F:1-245/Product: coat protein 1A #status predicted <P1A>  
F:246-491/Product: coat protein 1B #status predicted <P1B>  
F:492-836/Product: coat protein 1C #status predicted <P1C>  
F:837-980/Product: core protein 2A #status predicted <P2A>  
F:981-1076/Product: core protein 2B #status predicted <P2B>  
F:1077-1422/Product: core protein 2C #status predicted <P2C>

F:1423-1484/Product: protein 3A #status predicted <P3A>  
F:1453-1507/Product: protein 3B #status predicted <P3B>  
F:1508-1678/Product: cysteine proteinase, protein 3C #status predicted <P3C>  
F:1679-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 96.1%; Score 124; DB 1; Length 2227;  
Best Local Similarity 96.0%; Pred. No. 4.6e-10;

Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPKMKGLF 25  
|||||

DB 810 ORLKYAOEELSNEVLPPPKMKGLF 834

## RESULT 7

GNNYB

genome polypeptide - human hepatitis A virus (strain MB)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 1E; RNA-directed RNA polymerase (EC 2.7.7.48); protein 3D

Species: human hepatitis A virus  
Host: host Homo sapiens (man)

Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 10-May-1996  
Accession: J50303

R:Paul, A.V.; Tada, H.; von der Helm, K.; Wissel, T.; Klehn, R.; Wimmer, E.; Deinhardt, Virus Res. 8, 153-171, 1987

A:Title: The entire nucleotide sequence of the genome of human hepatitis A virus (isolat A:Reference number: J50303; MUID:88045071; PMID:2823500

A:Accession: J50303  
A:Molecule type: genomic RNA

A:Residues: 1-2227 <PAU>  
A:Cross-references: EMBL:M20273

C:Superfamily: hepatitis A virus genome polypeptide  
C:Keywords: coat protein; core protein; cysteine proteinase; genome-linked protein; hydrophobic

F:1-23/Product: coat protein 1A #status predicted <P4>  
F:24-246/Product: coat protein 1B #status predicted <VP2>

F:247-491/Product: coat protein 1C #status predicted <VP3>  
F:492-836/Product: coat protein 1D #status predicted <VP1>

F:837-980/Product: core protein 2A #status predicted <P2A>  
F:981-1108/Product: core protein 2B #status predicted <P2B>

F:1109-1438/Product: core protein 2C #status predicted <P2C>  
F:1439-1496/Product: core protein 2D #status predicted <P2D>

F:1497-1519/Product: genome-linked protein VPg #status predicted <VPg>  
F:1520-1736/Product: cysteine proteinase, protein 3C #status predicted <P3C>

F:1737-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <P3D>

Query Match 95.3%; Score 123; DB 1; Length 2227;  
Best Local Similarity 96.0%; Pred. No. 6.5e-10;

Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPKMKGLF 25  
|||||

DB 810 ORLKYAOEELSNEVLPPPKMKGLF 834

## RESULT 8

GNNYA

genome polypeptide - simian hepatitis A virus (strain AGM-27)

N:Contains: coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 1E; RNA-directed RNA polymerase (EC 2.7.7.48); protein 3D

Species: simian hepatitis A virus  
Host: host Homo sapiens (man)

Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 16-Jun-2000  
Accession: A30470; S03965

R:Tsarev, S.A.  
submitted to JIPID, April 1991

A:Reference number: A30470  
A:Accession: A30470  
A:Molecule type: genomic RNA

A:Residues: 1-2230 <TSA>  
A:Cross-references: GB:D00924; NID:9222597; PIDN:BA00766.1; PID:9222598

R:Tsarev, S.A.; Emerson, S.D.; Balayan, M.S.; Ticehurst, J.; Purcell, R.H.  
J. Gen. Virol. 72, 1677-1683, 1991  
A:Title: Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and  
A:Reference number: J01080; MUID:913111420; PMID:1649901  
A:Contents: annotation

A:Note: neither amino acid nor nucleotide sequence is given  
R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh  
Submitted to the EMBL Data Library, May 1989  
A:Reference number: S04885  
A:Accession: S04885

A:Molecule type: genomic RNA  
A:Residues: 1750-2164 <BAL1>

A:Cross-references: EMBL:X15461; NID:961971; PIDN:CAA3490.1; PID:9930268

R:Balayan, M.S.; Kusov, Y.Y.; Andjaparidze, A.G.; Tsarev, S.A.; Sverdlov, E.D.; Chizh  
FEBS Lett. 247, 425-428, 1989

A:Title: Variations in genome fragments coding for RNA polymerase in human and simian  
A:Reference number: S03965; MUID:89232168; PMID:2541023

A:Accession: S03965  
A:Molecule type: genomic RNA

A:Residues: 1960-2164 <BAL2>  
A:Cross-references: EMBL:X15461

C:Superfamily: hepatitis A virus genome polypeptide  
C:Keywords: coat protein; core protein; polypeptide

F:1-27/Product: coat protein 1A #status predicted <C1A>  
F:28-249/Product: coat protein 1B #status predicted <C1B>

F:250-495/Product: coat protein 1C #status predicted <C1C>  
F:496-795/Product: coat protein 1D #status predicted <C1D>

F:796-984/Product: core protein 2A #status predicted <C2A>  
F:985-1091/Product: core protein 2B #status predicted <C2B>

F:1092-1426/Product: core protein 2C #status predicted <C2C>  
F:1427-1498/Product: core protein 2D #status predicted <C2D>

F:1499-1521/Product: protein 3A #status predicted <P3A>  
F:1522-1741/Product: protein 3B #status predicted <P3B>

F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 92.2%; Score 119; DB 1; Length 2230;  
Best Local Similarity 88.0%; Pred. No. 2.6e-09;

Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPKMKGLF 25  
|||||

DB 814 ORLKYAOEELSNEVLPPPKMKGLF 838

## RESULT 9

GNNY2

genome polypeptide - simian hepatitis A virus (strain CY-145) (fragment)

N:Contains: amino end of core protein 2A; coat protein 1A; coat protein 1B; coat protein 1C; coat protein 1D; core protein 1E; RNA-directed RNA polymerase (EC 2.7.7.48); protein 3D

Species: simian hepatitis A virus  
Host: host Macaca fascicularis (cynomolgus macaque)

Date: 31-Mar-1993 #sequence\_revision 31-Mar-1993 #text\_change 16-Jul-1999  
Accession: J01180

R:Nathan, O.V.; Margolis, H.S.; Robertson, B.H.; Balayan, M.; Brinton, M.A.  
J. Gen. Virol. 72, 1685-1689, 1991

A:Title: Sequence analysis of a new hepatitis A virus naturally infecting cynomolgus  
A:Reference number: J01180; MUID:913111421; PMID:1649902

A:Accession: J01180  
A:Molecule type: genomic RNA

A:Residues: 1-839 <NAL>  
A:Cross-references: GB:M59286; NID:9329599; PIDN:AAA5473.1; PID:9555083

C:Superfamily: hepatitis A virus genome polypeptide  
C:Keywords: coat protein; core protein; glycoprotein; polypeptide

F:1-23/Product: coat protein 1A #status predicted <VP0>  
F:24-245/Product: coat protein 1B #status predicted <VP1>

F:246-491/Product: coat protein 1C #status predicted <VP2>  
F:492-839/Product: coat protein 2A (fragment) #status predicted <P2>

F:261,312,728,756/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 90.7%; Score 117; DB 1; Length 839;  
Best Local Similarity 84.0%; Pred. No. 1.7e-09;

Matches 21; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPKMKGLF 25  
|||||

DB 809 ORLKYAOEELSNEVLPPPKMKGLF 833

## RESULT 10

S74448  
regulatory protein pcr - *Synechocystis* sp. (strain PCC 6803)  
N:Alternate names: protein sll1408  
C:Species: *Synechocystis* sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 18-Aug-2000  
C:Accession: S74448  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp.  
S:  
A:Reference number: S74322; MUID:97061201; PMID:8905231  
A:Accession: S74448  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-346 <KAN>  
A:Cross-references: EMBL:D90899; GB:AB001339; NID:g1651650; PIDN:BAAL6600.1; PID:g165167  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
Gene: pcr  
Superfamily: *Synechocystis* regulatory protein pcr  
Keywords: transcription regulation

Query Match 43.8%; Score 56.5; DB 2; Length 346;  
Best Local Similarity 40.0%; Pred. No. 0.95;  
Matches 14; Conservative 4; Mismatches 4; Indels 13; Gaps 1;

OY 1 ORLKYAOEELSNEVLP 16  
||:|||||:|:|:|  
Db 241 ERKYAOEILYKVLPDPPLAQLSRQVSLNERKIK 275

RESULT 11  
T43701  
DNA-directed RNA polymerase (EC 2.7.7.6) II largest chain [Imported] - *Caenorhabditis* el  
C:Species: *Caenorhabditis* elegans  
C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 02-Sep-2000  
C:Accession: T43701  
R:Sidow, A.; Thomas, W.K.  
Curr. Biol. 4, 596-603, 1994  
A:Title: A molecular evolutionary framework for eukaryotic model organisms.  
A:Reference number: Z2636; MUID:95041334; PMID:7953533  
A:Accession: T43701  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-859 <SID>  
A:Cross-references: EMBL:U10333; NID:g520506; PIDN:AAA50224.1; PID:g520507  
A:Superfamily: DNA-directed RNA polymerase 132k polypeptide  
Keywords: nucleotidyltransferase

Query Match 40.3%; Score 52; DB 2; Length 859;  
Best Local Similarity 62.5%; Pred. No. 13;  
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 ORLKYAOEELSNEVLP 16  
||:|||||:|:|:|  
Db 141 ORIKYAREILQKELLP 156

RESULT 12  
E88445  
protein C26E.4 [Imported] - *Caenorhabditis* elegans  
C:Species: *Caenorhabditis* elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-May-2001  
C:Accession: E88445  
R:Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele  
A:Accession: E88445

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1193 <STO>  
A:Cross-references: GB:chr\_III; PIDN:AAA21158.1; PID:g532805; GSPDB:GN00021; CESP:C26  
C:Genetics:  
A:Gene: C26E.4  
A:Map position: 3  
C:Superfamily: DNA-directed RNA polymerase 132k polypeptide

Query Match 40.3%; Score 52; DB 2; Length 1193;  
Best Local Similarity 62.5%; Pred. No. 19;  
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 ORLKYAOEELSNEVLP 16  
||:|||||:|:|:|  
Db 340 ORIKYAREILQKELLP 355

RESULT 13  
T50995  
related to cytoskeleton assembly control protein SLA1 [Imported] - *Neurospora crassa*  
N:Alternate names: protein B7F18.140  
C:Species: *Neurospora crassa*  
C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000  
C:Accession: T50995  
R:Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu  
submitted to the Protein Sequence Database, July 2000  
A:Reference number: Z25286  
A:Accession: T50995  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1119 <SCH>  
A:Cross-references: EMBL:AL389891; GSPDB:GN00116; NCSP:B7F18.140  
A:Experimental source: BAC clone B7F18; strain OR74A  
C:Genetics:  
A:Gene: NCSP:B7F18.140  
A:Map position: 6  
A:Insertions: 66/3; 123/2; 495/1

Query Match 39.5%; Score 51; DB 2; Length 1119;  
Best Local Similarity 55.6%; Pred. No. 25;  
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

OY 1 ORLKYAOEELSNEVLP 18  
||:|||||:|:|:|  
Db 165 QROSYASEDYENYRSP 182

RESULT 14  
T29757  
protein UNC-89 - *Caenorhabditis* elegans  
C:Species: *Caenorhabditis* elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 03-Dec-1999  
C:Accession: T29757  
R:Du, Z.; Le, T.T.; Wilson, R.  
submitted to the EMBL Data Library, May 1997  
A:Description: The sequence of *C. elegans* cosmid C09D1.  
A:Reference number: Z20679  
A:Accession: T29757  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-6642 <DUZ>  
A:Cross-references: EMBL:AF003131; PIDN:AAB54132.1; GSPDB:GN00019; CESP:unc-89  
A:Experimental source: strain Bristol N2; clone C09D1  
C:Genetics:  
A:Gene: CESP:unc-89  
A:Map position: 1  
A:Insertions: 17/2; 108/3; 154/2; 211/2; 265/3; 326/2; 352/3; 426/2; 454/1; 500/1; 537/1  
/3; 5917/1; 6027/1; 6061/3; 6153/2; 6515/1; 6553/3; 6609/1

Query Match 39.5%; Score 51; DB 2; Length 6642;  
Best Local Similarity 45.5%; Pred. No. 1.9e+02;  
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 ; Search time 8.23529 Seconds

(without alignments)  
142.760 Million cell updates/sec

Title: US-09-171-432a-47

Perfect score: 129

Sequence: 1 QRLRYAOEELSNEVLPPEPKMKGLF 25

Scoring table: BLOSUM62

Searched: 127863 seqs, 47026705 residues

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	341	1	P13672 hepatitis a
2	129	100.0	852	1	P06442 hepatitis a
3	129	100.0	2226	1	P06580 hepatitis a
4	129	100.0	2226	1	P26581 hepatitis a
5	129	100.0	2226	1	P26582 hepatitis a
6	129	100.0	2227	1	P06617 hepatitis a
7	129	100.0	2227	1	P06641 hepatitis a
8	123	95.3	2227	1	P13901 hepatitis a
9	119	92.2	2230	1	P14553 simian hepa
10	117	90.7	839	1	P03178 simian hepa
11	103	79.8	808	1	P02381 hepatitis a
12	53	41.1	636	1	Q15025 homo sapien
13	52	40.3	1193	1	Q10578 caenorhabdi
14	51	39.5	6632	1	Q01761 caenorhabdi
15	50.5	39.1	443	1	Q9K977 vibrio chol
16	49.5	38.4	554	1	O8dcp4 vibrio vuln
17	49	38.0	1174	1	P51021 podospira a
18	48.5	37.6	2278	1	P30876 homo sapien
19	48	37.2	1176	1	P34756 saccharomyc
20	48	37.2	1191	1	P08266 drosophila
21	48	37.2	1210	1	Q42872 lycopersico
22	48	36.4	578	1	Q02061 schizosacch
23	47	36.4	592	1	P09765 influenza b
24	47	36.0	399	1	P15891 saccharomyc
25	46.5	35.7	338	1	P34821 mus musculu
26	46	35.7	480	1	Q01664 mus sapien
27	46	35.7	574	1	P13744 cucurbita m
28	46	35.7	576	1	P09766 influenza b
29	46	35.7	578	1	P09767 influenza b
30	46	35.7	583	1	P10757 influenza b
31	46	35.7	583	1	P03464 influenza b
32	46	35.7	583	1	P03464 influenza b
33	46	35.7	583	1	P03463 influenza b

34	46	35.7	585	1	HEMA_INBE	P17504 influenza b
35	46	35.7	585	1	HEMA_INBYK	P22092 influenza b
36	46	35.7	1085	1	CARB_HELPJ	P09472 helicobacte
37	45	34.9	315	1	RSPB_HAEIN	P44792 haemophilus
38	45	34.9	445	1	ML64_HUMAN	Q14849 homo sapien
39	45	34.9	733	1	HEXA_BLD1	Q17127 blaberus di
40	45	34.9	1188	1	RPB2_ARATH	P38420 arabidopsis
41	45	34.9	1312	1	L159_CAEEL	P17886 drosophila
42	44.5	34.5	702	1	CRR_DROME	P17886 drosophila
43	44.5	34.5	3678	1	DMD_MOUSE	P11531 mus musculu
44	44.5	34.5	3680	1	DMD_CANFA	O97592 canis fam11
45	44.5	34.5	3685	1	DMD_HUMAN	P11532 homo sapien

## ALIGNMENTS

```

RESULT 1
POLG_HP1AV1 STANDARD; PRT; 341 AA.
AC P13672:
DT 01-JAN-1990 (rel. 13, Created)
DT 01-JAN-1990 (rel. 13, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP3; Core protein
DE P2A] (Fragment).
OS Hepatitis A virus (strain CDC-1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OC NCBI_TaxID=12093;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89263805; PubMed=2542903;
RA Andonov A.P., Lau P., Chaudhary R.;
RT "Nucleotide sequence of the VP1 gene from a Chinese strain of
RT hepatitis A virus (HAV).";
RL Nucleic Acids Res. 17:3594-3594(1989).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VITRO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X14666; CA32794.1; -.
DR PIR: S04137; S04137.
KW Polyprotein; Coat protein; Core protein.
FT NON_TER 1
FT CHAIN <1 1 COAT PROTEIN VP3 (1C).
FT CHAIN 2 340 COAT PROTEIN VP1 (1D).
FT CHAIN 341 >341 CORE PROTEIN P2A.
FT NON_TER 341
SQ SEQUENCE 341 AA; 38003 MW; 066918289BF126D5 CRC64;

Query Match 100.0%; Score 129; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 6,2e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QRLRYAOEELSNEVLPPEPKMKGLF 25
DB 314 QRLRYAOEELSNEVLPPEPKMKGLF 338

RESULT 2
POLG_HP1AV1 STANDARD; PRT; 852 AA.
AC P06442; O83741; O83742;

```

DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core protein P2A] (Fragment).  
OS Hepatitis A virus (strain CR326).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=8185648; PubMed=2985793;  
RA Linemeyer D.L., Menke J.G., Martin-Gallardo A., Hughes J.V.,  
RT Young A., Mitra S.W.,  
RL "Molecular cloning and partial sequencing of hepatitis A viral cDNA."  
J. Virol. 54:247-255(1985).  
-1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,  
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
VP3, AND VP4.  
-1- PPM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: M10033; AAA45470.1; -.  
DR PIR: A03904; GNNYHA.  
KW Polypeptide; Coat protein; Core protein.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 836  
FT CHAIN 837 >852  
FT NON\_TER 852 852  
SQ SEQUENCE 852 AA; 95563 MW; 73D3ED0AD53820E CRC64;  
  
Query Match 100.0%; Score 129; DB 1; Length 852;  
Best Local Similarity 100.0%; Pred. No. 1.7e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QRLKYAOEELSENYLPPRRKMKGLF 25  
DB 810 QRLKYAOEELSENYLPPRRKMKGLF 834  
-----  
ID POLG\_HPAV2 STANDARD; PRT; 2226 AA.  
AC P26580;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins  
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
DE P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain 24a).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12094;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91162758; PubMed=1705995;  
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,  
RT Cromeans T., Jansen R.W.,  
RT "Antigenic and genetic variation in cytopathic hepatitis A virus  
RT variants arising during persistent infection: evidence for genetic  
RT recombination."  
J. Virol. 65:2056-2065(1991).

CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC (RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PPM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: M59810; AAA45468.1; -.  
DR MEROPS: C03.005; -.  
DR InterPro: IPR004004; CalicI\_Pol\_hel.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR007094; RNA\_pol\_PSVir.  
DR pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
DR pfam: PF00910; RNA\_helicase; 1.  
DR PRINTS: PR00918; CALICIVIRUSNS.  
KW RNA-directed RNA polymerase; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 794  
FT CHAIN 795 900  
FT CHAIN 901 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1495  
FT CHAIN 1496 1518  
FT CHAIN 1519 1737  
FT CHAIN 1738 2226  
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91D6B4E2BF CRC64;  
  
Query Match 100.0%; Score 129; DB 1; Length 2226;  
Best Local Similarity 100.0%; Pred. No. 4.9e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 QRLKYAOEELSENYLPPRRKMKGLF 25  
DB 810 QRLKYAOEELSENYLPPRRKMKGLF 834  
-----  
ID POLG\_HPAV4 STANDARD; PRT; 2226 AA.  
AC P26581;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins  
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
DE P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain 43c).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12095;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91162758; PubMed=1705995;  
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,  
RT Cromeans T., Jansen R.W.,  
RT "Antigenic and genetic variation in cytopathic hepatitis A virus  
RT variants arising during persistent infection: evidence for genetic  
RT recombination."  
J. Virol. 65:2056-2065(1991).

```
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59809: AAA45469.1; -.
CC
CC DR MEROPS: C03.005; -.
CC DR InterPro: IPR000605; RNA_helicase.
CC DR InterPro: IPR007095; RNA_pol_DS_PS.
CC DR InterPro: IPR001205; RNA_pol_P3D.
CC DR InterPro: IPR007094; RNA_pol_PSVlr.
CC DR Pfam: PF00680; RNA_dep_RNA_pol.1.
CC
CC DR Pfam: PF00910; RNA_helicase.1.
CC DR polyprotein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease.
CC
CC FT CHAIN 1 23
CC FT CHAIN 24 245
CC FT CHAIN 246 491
CC FT CHAIN 492 794
CC FT CHAIN 795 900
CC FT CHAIN 901 1087
CC FT CHAIN 1088 1422
CC FT CHAIN 1423 1495
CC FT CHAIN 1496 1518
CC FT CHAIN 1519 1737
CC FT CHAIN 1738 2226
CC FT CHAIN 2226 251107 MW; 403B4CA80B09BE75 CRC64;
CC
CC SQ SEQUENCE
CC
CC Query Match 100.0%; Score 129; DB 1; Length 2226;
CC Best Local Similarity 100.0%; Pred. No. 4.9e-11;
CC Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 QRLKYAOEELSNEVLPPrKKMKGLF 25
CC DB 810 QRLKYAOEELSNEVLPPrKKMKGLF 834
CC
CC RESULT 5
CC POLG_HPAV8 STANDARD: PRT; 2226 AA.
CC ID POLG_HPAV8
CC AC P26582;
CC DT 01-AUG-1992 (Rel. 23, Created)
CC DT 01-AUG-1992 (Rel. 23, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
CC P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
CC P3D (EC 2.7.7.48)].
CC OS Hepatitis A virus (strain 18f).
CC OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
CC Hepatovirus.
CC NCBI_TaxID=12096;
CC OX NCBI_TaxID=12096;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=91162758; PubMed=1705995;
CC RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstein S.M.,
CC Cromeans T., Jansen R.W.;
CC RT Antigenic and genetic variation in cytopathic hepatitis A virus
CC RT variants arising during persistent infection: evidence for genetic
CC RT recombination.
CC RL J. Virol. 65:2056-2065(1991).
CC CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC [RNA](N).
```

```
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M59808: AAA45467.1; -.
CC
CC DR PDB: 1OAT; 15-MAY-00.
CC
CC DR MEROPS: C03.005; -.
CC DR InterPro: IPR000605; RNA_helicase.
CC DR InterPro: IPR007095; RNA_pol_DS_PS.
CC DR InterPro: IPR001205; RNA_pol_P3D.
CC DR InterPro: IPR007094; RNA_pol_PSVlr.
CC DR Pfam: PF00680; RNA_dep_RNA_pol.1.
CC
CC DR Pfam: PF00910; RNA_helicase.1.
CC DR polyprotein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease; 3D-structure.
CC
CC FT CHAIN 1 23
CC FT CHAIN 24 245
CC FT CHAIN 246 491
CC FT CHAIN 492 794
CC FT CHAIN 795 900
CC FT CHAIN 901 1087
CC FT CHAIN 1088 1422
CC FT CHAIN 1423 1495
CC FT CHAIN 1496 1518
CC FT CHAIN 1519 1737
CC FT CHAIN 1738 2226
CC FT CHAIN 2226 251292 MW; 24964A6396C8D6B CRC64;
CC
CC SQ SEQUENCE
CC
CC Query Match 100.0%; Score 129; DB 1; Length 2226;
CC Best Local Similarity 100.0%; Pred. No. 4.9e-11;
CC Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 QRLKYAOEELSNEVLPPrKKMKGLF 25
CC DB 810 QRLKYAOEELSNEVLPPrKKMKGLF 834
CC
CC RESULT 6
CC POLG_HPAV8 STANDARD: PRT; 2227 AA.
CC ID POLG_HPAV8
CC AC P08617; P06443; O81082;
CC DT 01-AUG-1988 (Rel. 08, Created)
CC DT 01-AUG-1988 (Rel. 08, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins
CC P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
CC P3D (EC 2.7.7.48)].
CC OS Hepatitis A virus (strain HM-175).
CC OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
CC Hepatovirus.
CC NCBI_TaxID=12098;
CC OX NCBI_TaxID=12098;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RX STRAIN=Wild type;
CC RX MEDLINE=87061253; PubMed=3023706;
CC RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
CC Baroudy B.M.;
CC RT Complete nucleotide sequence of wild-type hepatitis A virus:
CC RT comparison with different strains of hepatitis A virus and other
CC RT picornaviruses.
CC RL J. Virol. 61:50-59(1987).
CC [2]
CC RN [2]
CC RP SEQUENCE FROM N.A.
```

RC STRAIN-Attenuated;  
 RA MEDLINE-87175701; PubMed-3031686;  
 RA Cohen J.I., Rosenblum B., Ticehurst J.R., Daemer R.J., Felstone S.M.,  
 RA Purcell R.H.;  
 RT "Complete nucleotide sequence of an attenuated hepatitis A virus:  
 RT comparison with wild-type virus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).  
 RN [1]  
 RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.  
 RX MEDLINE-85166289; PubMed-2984684;  
 RA Baroudy B.M., Ticehurst J.R., Miele T.A., Matzel J.V. Jr.,  
 RA Purcell R.H., Felstone S.M.;  
 RT "Sequence analysis of hepatitis A virus cDNA coding for capsid  
 RT proteins and RNA polymerase.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +  
 CC (RNA)(N).  
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS.  
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
 CC VP3, AND VP4.  
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC -1- MISCELLANEOUS: STRAIN HM-175/7 MK-5 IS ATTENUATED STRAIN DERIVED  
 CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT  
 CC SHOWN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M14114; AAA45475.1; -;  
 CC EMBL: M14707; AAA45465.1; -;  
 CC EMBL: M16632; AAA45471.1; -;  
 CC PIR: A25981; GNNYHM.  
 CC PIR: A94149; GNNYMK.  
 CC PDB: 1HAV; 23-DEC-96.  
 CC MEROPS: C03.005; -;  
 CC InterPro: IPR004004; Callc1\_pol\_hel.  
 CC InterPro: IPR000605; RNA\_helicase.  
 CC InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 CC InterPro: IPR001205; RNA\_pol\_P3D.  
 CC InterPro: IPR007094; RNA\_pol\_PSVlr.  
 CC Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
 CC Pfam: PF00910; RNA\_helicase; 1.  
 CC PRINTS: PR00918; CALICIVIRUSNS.  
 CC -----  
 CC Polypeptide: Coat protein; Core protein; Transferase;  
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.  
 CC -----  
 CC CHAIN 1 23  
 CC FT CHAIN COAT PROTEIN VP4 (PIA).  
 CC FT CHAIN COAT PROTEIN VP2 (PIB).  
 CC FT CHAIN COAT PROTEIN VP3 (PIC).  
 CC FT CHAIN COAT PROTEIN VP1 (PID).  
 CC FT CHAIN CORE PROTEIN P2A.  
 CC FT CHAIN CORE PROTEIN P2B.  
 CC FT CHAIN CORE PROTEIN P2C.  
 CC FT CHAIN PROBABLE PROTEIN P3A.  
 CC FT CHAIN PROBABLE PROTEIN P3B.  
 CC FT CHAIN PROBABLE PROTEIN P3C.  
 CC FT CHAIN RNA-DIRECTED POLYMERASE 3D.  
 CC FT CHAIN K -> R (IN ATTENUATED STRAIN).  
 CC FT CHAIN E -> V (IN ATTENUATED STRAIN).  
 CC FT CHAIN N -> S (IN ATTENUATED STRAIN).  
 CC FT CHAIN A -> V (IN ATTENUATED STRAIN).  
 CC FT CHAIN G -> A (IN ATTENUATED STRAIN).  
 CC FT CHAIN K -> M (IN ATTENUATED STRAIN).  
 CC FT CHAIN E -> K (IN ATTENUATED STRAIN).  
 CC FT CHAIN F -> S (IN ATTENUATED STRAIN).  
 CC FT CHAIN V -> I (IN ATTENUATED STRAIN).  
 CC FT CHAIN 1277 1277

FT VARIANT 1500 1500 H -> Y (IN ATTENUATED STRAIN).  
 FT VARIANT 1805 1805 D -> N (IN ATTENUATED STRAIN).  
 FT VARIANT 1930 1930 S -> T (IN ATTENUATED STRAIN).  
 SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AEB740A6 CRC64;  
 Query Match 100.0%; Score 129; DB 1; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 4; 9e-11;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 ORLKYAOEISNEVLPPEPRKMKGLF 25  
 Db 810 ORLKYAOEISNEVLPPEPRKMKGLF 834  
 RESULT 7  
 ID POLG\_HPAVL STANDARD; PRT; 2227 AA.  
 AC P06441;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polypeptide [contains: Coat proteins VP1 TO VP4; Core proteins  
 DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
 DE P3D (PC 2.7.7.48)].  
 OS Hepatitis A virus (strain Ia).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Hepatovirus.  
 OX NCBI\_TaxID=12099;  
 RX SEQUENCE FROM N.A.  
 RX MEDLINE-85190549; PubMed-2986127;  
 RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,  
 RA Merryweather J., van Nest G., Dina D.;  
 RT "Primary structure and gene organization of human hepatitis A virus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +  
 CC (RNA)(N).  
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,  
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
 CC VP3, AND VP4.  
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: K02990; AAA45472.1; -;  
 CC PIR: A03903; GNNYHR.  
 CC MEROPS: C03.005; -;  
 CC InterPro: IPR004004; Callc1\_pol\_hel.  
 CC InterPro: IPR000605; RNA\_helicase.  
 CC InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 CC InterPro: IPR001205; RNA\_pol\_P3D.  
 CC InterPro: IPR007094; RNA\_pol\_PSVlr.  
 CC Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
 CC Pfam: PF00910; RNA\_helicase; 1.  
 CC PRINTS: PR00918; CALICIVIRUSNS.  
 CC -----  
 CC Polypeptide: Coat protein; Core protein; Transferase;  
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
 CC -----  
 CC CHAIN 1 23  
 CC FT CHAIN COAT PROTEIN VP4 (PIA).  
 CC FT CHAIN COAT PROTEIN VP2 (PIB).  
 CC FT CHAIN COAT PROTEIN VP3 (PIC).  
 CC FT CHAIN COAT PROTEIN VP1 (PID).  
 CC FT CHAIN CORE PROTEIN P2A.  
 CC FT CHAIN CORE PROTEIN P2B.  
 CC FT CHAIN CORE PROTEIN P2C.  
 CC FT CHAIN PROBABLE PROTEIN P3A.  
 CC FT CHAIN PROBABLE PROTEIN P3B.  
 CC FT CHAIN 1485 1507

FT CHAIN 1508 1678 PROBABLE PROTEIN P3C.  
 FT CHAIN 1679 2227 RNA-DIRECTED POLYMERASE P3D.  
 SQ SEQUENCE 2227 AA: 251898 MW: 99A7354B4CD2799C CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 4,9e-11;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORLKYAOEELSNEVLPPPRKMGFLF 25  
 DB 810 ORLKYAOEELSNEVLPPPRKMGFLF 834

## RESULT 8

POLG\_HPAVM STANDARD; PRT: 2227 AA.  
 ID POLG\_HPAVM 081083; 081084; 081085; 081086; 081087; 081088; 081089;  
 AC 081090; 081091; 081092; 081093;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DE 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
 DE Hepatitis A virus (strain MBB).  
 OS Hepatitis A virus (strain MBB).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Hepatovirus.  
 OX NCBI\_TaxID=12100;

RP SEQUENCE FROM N.A.  
 RX MEDLINE=88045071; PubMed=2823500;  
 RA Paul A.V., Tada H., der Helm K., Wessel T., Klein R., Wimmer E.,  
 RA Delnhardt F.;  
 RT "The entire nucleotide sequence of the genome of human hepatitis A virus (isolate MBB).";  
 RL Virus Res. 8:153-171(1987).  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate + {RNA}(N).

CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.  
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC -1- MISCELLANEOUS: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.

CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M20273; AAA5474.1; -  
 DR MEROPS: C03.005; -  
 DR InterPro: IPR000605; RNA\_helicase.  
 DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro: IPR001205; RNA\_pol\_P3D.  
 DR InterPro: IPR007094; RNA\_pol\_PSVIR.  
 DR Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam: PF00910; RNA\_helicase; 1.  
 KW Polyprotein; Coat protein; Core protein; Transferase;  
 KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.

FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).  
 FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).  
 FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).  
 FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).  
 FT CHAIN 837 980 CORE PROTEIN P2A.  
 FT CHAIN 981 1087 CORE PROTEIN P2B.  
 FT CHAIN 1088 1432 CORE PROTEIN P2C.  
 FT CHAIN 1433 1496 PROBABLE PROTEIN P3A.  
 FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.

FT CHAIN 1520 1738 PROBABLE PROTEIN P3C.  
 FT CHAIN 1739 2227 RNA-DIRECTED POLYMERASE 3D.  
 SQ SEQUENCE 2227 AA: 251425 MW: EC983EDDA7C86349 CRC64;

Query Match 95.3%; Score 123; DB 1; Length 2227;  
 Best Local Similarity 96.0%; Pred. No. 3,9e-10;  
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ORLKYAOEELSNEVLPPPRKMGFLF 25  
 DB 810 ORLKYAOEELSNEVLPPPRKMGFLF 834

## RESULT 9

POLG\_HPAVS STANDARD; PRT: 2230 AA.  
 ID POLG\_HPAVS 014553;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
 DE Simian hepatitis A virus (strain AGM-27).  
 OS Simian hepatitis A virus (strain AGM-27).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Hepatovirus.  
 OX NCBI\_TaxID=12102;

RP SEQUENCE FROM N.A.  
 RX MEDLINE=91311420; PubMed=1649901;  
 RA Tsarev S.A., Emerson S.U., Balayan M.S., Ticehurst J.R.,  
 RA Purcell R.H.;  
 RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome structure and growth in cell culture with other HAV strains.";  
 RL J. Gen. Virol. 72:1677-1683(1991).  
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.

CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: D00924; BAA00766.1; -  
 DR EMBL: X15461; CAA33490.1; -  
 DR PIR: A30470; GNNYSA.  
 DR MEROPS: C03.005; -  
 DR InterPro: IPR004004; Calic1\_pol\_hel.  
 DR InterPro: IPR000605; RNA\_helicase.  
 DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro: IPR001205; RNA\_pol\_P3D.  
 DR InterPro: IPR007094; RNA\_pol\_PSVIR.  
 DR Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam: PF00910; RNA\_helicase; 1.  
 KW PRINTS: PR00918; CALICVIRUSNS.  
 KW Polyprotein; Coat protein; Core protein; Transferase;  
 KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease.

FT CHAIN 1 23 COAT PROTEIN VP4 (P1A).  
 FT CHAIN 24 245 COAT PROTEIN VP2 (P1B).  
 FT CHAIN 246 491 COAT PROTEIN VP3 (P1C).  
 FT CHAIN 492 836 COAT PROTEIN VP1 (P1D).  
 FT CHAIN 837 980 CORE PROTEIN P2A.  
 FT CHAIN 981 1087 CORE PROTEIN P2B.  
 FT CHAIN 1088 1432 CORE PROTEIN P2C.  
 FT CHAIN 1433 1496 PROBABLE PROTEIN P3A.  
 FT CHAIN 1497 1519 PROBABLE PROTEIN P3B.

```
FT CHAIN 1 27 COAT PROTEIN VP4 (PIA).
FT CHAIN 28 249 COAT PROTEIN VP2 (PIB).
FT CHAIN 250 495 COAT PROTEIN VP3 (PIC).
FT CHAIN 496 795 COAT PROTEIN VP1 (PID).
FT CHAIN 796 984 CORE PROTEIN P2A.
FT CHAIN 985 1091 CORE PROTEIN P2B.
FT CHAIN 1092 1426 CORE PROTEIN P2C.
FT CHAIN 1427 1498 PROBABLE PROTEIN 3A.
FT CHAIN 1499 1521 PROBABLE PROTEIN 3B.
FT CHAIN 1522 1741 PROBABLE PROTEIN 3C.
FT CHAIN 1742 2230 RNA-DIRECTED POLYMERASE 3D.
SO SEQUENCE 2230 AA; 251296 MW; 8783230E324E1F19 CRC64;

Query Match 92.2%; Score 119; DB 1; Length 2230;
Best Local Similarity 88.0%; Pred. No. 1.6e-09;
Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPKKGLF 25
||||| ||||||| ||||||| |||
814 ORLKYAOEELSNEVLPPPKKGLF 838

POLG_HPAVT STANDARD; PRT; 839 AA.
ID P31768;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 to VP4; Core protein
P2A] (Fragment).
OS Simian hepatitis A virus (strain CY-145).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
Hepatovirus.
NCBI_TaxID=31707;
RN [1]
RP MEDLINE=91311421; PubMed=1649902;
RA Nainan O.V., Margolis H.S., Robertson B.H., Brlinton M.A.;
RT "Sequence analysis of a new hepatitis A virus naturally infecting
cytomegalus macaques (Macaca fascicularis).";
RL J. Gen. Virol. 72:1685-1689(1991).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M59286; AAA5473.1; -.
DR PIR: J01180; GNMKX2.
KW Polyprotein; Coat protein; Core protein.
FT CHAIN 1 23 COAT PROTEIN VP4 (PIA).
FT CHAIN 24 245 COAT PROTEIN VP2 (PIB).
FT CHAIN 246 491 COAT PROTEIN VP3 (PIC).
FT CHAIN 492 ? COAT PROTEIN VP1 (PID).
FT CHAIN ? >839 CORE PROTEIN P2A.
FT NON_TER 839 839
SO SEQUENCE 839 AA; 93825 MW; 2CACC4BD1E192DBC CRC64;

Query Match 90.7%; Score 117; DB 1; Length 839;
Best Local Similarity 84.0%; Pred. No. 1.1e-09;
Matches 21; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPKKGLF 25
||||| ||||||| ||||||| |||
809 ORLKYAOEELSNEVLPPPKKGLF 833
```

```
RESULT 11
POLG_HPAVT STANDARD; PRT; 808 AA.
ID Q02381;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 to VP4; Core protein
P2A] (Fragment).
OS Hepatitis A virus (strain GA/76).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
Hepatovirus.
NCBI_TaxID=31706;
RN [1]
RP MEDLINE=92260183; PubMed=1316423;
RA Khanna B., Speldring J.E., Innis B.L., Robertson B.H.;
RT "Characterization of a genetic variant of human hepatitis A virus.";
RL J. Med. Virol. 36:118-124(1992).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,
EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M66695; AAA5477.1; -.
KW Polyprotein; Coat protein; Core protein.
FT NON_TER 1 1
FT CHAIN 1 2 COAT PROTEIN VP4 (PIA).
FT CHAIN 3 223 COAT PROTEIN VP2 (PIB).
FT CHAIN 224 470 COAT PROTEIN VP3 (PIC).
FT CHAIN 471 770 COAT PROTEIN VP1 (PID).
FT CHAIN 771 >808 CORE PROTEIN P2A.
FT NON_TER 808 808
SO SEQUENCE 808 AA; 90632 MW; D80CE7E57A479C12 CRC64;

Query Match 79.8%; Score 103; DB 1; Length 808;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ORLKYAOEELSNEVLPPPK 20
||||| ||||||| ||||||| |||
789 ORLKYAOEELSNEVLPPPK 808

RESULT 12
NAFL_HUMAN STANDARD; PRT; 636 AA.
ID Q15025; Q96008; Q96EL9; Q99833; Q9HJ3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Nef-associated factor 1 (Naf1) (HIV-1 Nef interacting protein)
DE (Viron-associated nuclear shuttling protein) (VAN) (hVAN) (TNPAIP3
interacting protein 1).
GN TNIP1 OR NAF1 OR KIA0113.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Peripheral blood;
RX MEDLINE=99120485; PubMed=9923610;
```

RA Yamamoto N.;  
RA Fukushima M., Dixon J., Kimura T., Tsurutani N., Dixon M.J.,  
RT "Identification and cloning of a novel cellular protein Nafl, Net-  
RT associated factor 1, that increases cell surface CD4 expression";  
RL FEMS Lett. 442:83-88(1999).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Leukocyte;  
RX MEDLINE=20541981; PubMed=11090181;  
RA Gupta K., Ott D., Hope T.J., Silliciano R.F., Boeke J.D.;  
RT "A human nuclear shuttling protein that interacts with human  
RT immunodeficiency virus type 1 matrix is packaged into virions";  
RL J. Virol. 74:11811-11824(2000).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Lung;  
RX MEDLINE=22386257; PubMed=12477932;  
RA Strausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
ALtschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
Stempkin M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
Brenstein M.J., Udell T.B., Toshiyuki S., Cantinici P., Prange C.,  
Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
Bosak S.A., McEwan P.C., McKernan K.J., Malek J.A., Gunaratne P.H.,  
Richards S., Wotley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
Falley J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M.,  
Butcherfield Y.S.N., Krzyzanski M.I., Skalska U., Smellus D.E.,  
Sutnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RA "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).  
RN [4]  
RP SEQUENCE OF 136-636 FROM N.A. (ISOFORM 2).  
RC TISSUE=Craniofacial;  
RX MEDLINE=96276047; PubMed=8681136;  
RA Loftus S.K., Dixon J., Koprivnikar K., Dixon M.J., Wasmuth J.J.;  
RT "Genomic map of the Treacher Collins candidate gene region";  
RL Genomics Res. 6:26-34(1996).  
RN [5]  
RP SEQUENCE OF 341-636 FROM N.A. (ISOFORM 1).  
RC TISSUE=Bone marrow;  
RX MEDLINE=95308325; PubMed=7786527;  
RA Nagase T., Miyajima N., Tanaka A., Sazuka T., Seki N., Sato S.,  
Tabata S., Ishikawa K.-I., Kawarabayashi Y., Kotani H., Nomura N.;  
RT "Prediction of the coding sequences of unidentified human genes. III:  
RT The coding sequences of 40 new genes (K1A00081-K1A0120) deduced by  
RL analysis of cDNA clones from human cell line KG-1";  
RN DNA Res. 2:37-43(1995).  
RN [6]  
RP SEQUENCE OF 94-412 FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=95308325; PubMed=7786527;  
RA Fukushima M., Kimura T., Yamamoto N.;  
RL Submitted (Jan-1997) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP FUNCTION: Interacts with zinc finger protein A20/TNFAIP3 and  
RC inhibits TNF-induced NF-kappa-B-dependent gene expression by  
RX interfering with an RIP- or TRAF2-mediated transactivation signal  
RN (by similarity). Increases cell surface CD4(74) antigen  
CC expression. Interacts with HIV-1 matrix protein and is packaged  
CC into virions and overexpression can inhibit viral replication. May  
CC regulate matrix nuclear localization, both nuclear import of PIC  
CC (preintegration complex) and export of Gag polypeptide and viral  
CC genomic RNA during virion production.  
CC [8]  
CC SUBUNIT: Interacts with TNFAIP3 (By similarity). Interacts with  
CC HIV-1 matrix protein.  
CC [9]  
CC SUBCELLULAR LOCATION: Cytoplasmic. Shuttles between the nucleus  
CC and cytoplasm in a CRM1-dependent manner.  
CC [10]  
CC ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;

RP SEQUENCE OF 200-1058 FROM N.A.  
RX MEDLINE-95041334; PubMed-7953533;  
RA Sidow A., Thomas W.K.,  
RT "A molecular evolutionary framework for eukaryotic model organisms."  
RL Curr. Biol. 4:596-603(1994).  
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
CC SUBSTRATES.  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +  
CC (RNA)(N).  
CC -1- SUBUNIT: EACH CLASS OF RNA POLYMERASE IS ASSEMBLED FROM 9 TO 14  
CC DIFFERENT POLYPEPTIDES. THIS SUBUNIT IS THE SECOND LARGEST  
CC COMPONENT OF RNA POLYMERASE II.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- MISCELLANEOUS: THREE DISTINCT ZINC-CONTAINING RNA POLYMERASES ARE  
CC FOUND IN EUKARYOTIC NUCLEI: POLYMERASE I FOR THE RIBOSOMAL RNA  
CC PRECURSOR, POLYMERASE II FOR THE MRNA PRECURSOR, AND POLYMERASE  
CC III FOR 5S AND TRNA GENES.  
CC -1- SIMILARITY: Belongs to the RNA polymerase beta chain family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U13875; AAA21158.1; -  
CC EMBL; U10333; AAA50224.1; -  
CC PIR; E88445; E88445.  
CC PIR; T43701; T43701.  
CC WormPep; C26E6.4; CE01162.  
CC InterPro; IPR001572; RNA\_POL\_B.  
CC DR Pfam; PF04563; RNA\_POL\_Rpb2\_1; 1.  
CC DR Pfam; PF04561; RNA\_POL\_Rpb2\_2; 1.  
CC DR Pfam; PF04565; RNA\_POL\_Rpb2\_3; 1.  
CC DR Pfam; PF04566; RNA\_POL\_Rpb2\_4; 1.  
CC DR Pfam; PF04567; RNA\_POL\_Rpb2\_5; 1.  
CC DR Pfam; PF00562; RNA\_POL\_Rpb2\_6; 1.  
CC DR Pfam; PF04560; RNA\_POL\_Rpb2\_7; 1.  
CC DR Pfam; PF04560; RNA\_POL\_BETA; 1.  
CC DR PROSITE; PS01166; RNA\_POL\_POLYMERASE; Transcription; Zinc;  
CC KW Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;  
CC zinc-finger; Nuclear protein.  
CC FT ZN.FING 1125 1146 C4-TYPE (POTENTIAL).  
CC SQ SEQUENCE 1193 AA; 134904 MW; BBA05E74E9C7BDE CRC64;  
CC  
CC Query Match 40.3%; Score 52; DB 1; Length 1193;  
CC Best Local Similarity 62.5%; Pred. No. 9.3;  
CC Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0.  
CC  
CC Oy 1 ORLKYAOBELSNEVLP 16  
CC ||:||||:| |||  
CC Db 340 QRIKYARELLQKELLP 355  
CC  
CC RESULT 14  
CC UN89\_CAEEL  
CC ID UN89\_CAEEL STANDARD; PRT; 6632 AA.  
CC AC 001761; Q17362;  
CC DT 15-SEP-2003 (Rel. 42, Created)  
CC DT 15-SEP-2003 (Rel. 42, Last sequence update)  
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)  
CC DE Muscle M-line assembly protein unc-89 (Uncoordinated protein 89).  
CC UNC-89 OR C09D1.1.  
CC Caenorhabditis elegans.  
CC OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;  
CC Rhabditidae; Peloderinae; Caenorhabditis.  
CC NCBI\_TaxID=62339;  
CC [1]  
CC RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.  
CC RC STRAIN-Bristol N2;  
CC RX MEDLINE-96180278; PubMed-8603916;

RA	Benian G.M., Tinley T.L., Tang X., Borodovsky M.;		
RA	"The Caenorhabditis elegans gene unc-89, required for muscle M-line		
RT	assembly, encodes a giant modular protein composed of Ig and signal		
RT	transduction domains."		
RL	J. Cell Biol. 132:835-848(1996).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-Bristol N2;		
RA	Du Z., Le T.T., Wilson R.;		
RL	Submitted (May-1997) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	REVIEWS.		
RA	Waterston R.;		
CC	Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.		
CC	-1- FUNCTION: Structural component of the muscle M-line. Myofibril		
CC	lattice assembly begins with positional cues laid down in the		
CC	basement membrane and muscle cell membrane. UNC-89 responds to		
CC	these signals, localizes, and then participates in assembling an		
CC	M-line.		
CC	-1- SIMILARITY: Localizes to the middle of A-bands.		
CC	-1- SIMILARITY: Contains 1 DBJ-homology (DH) domain.		
CC	-1- SIMILARITY: Contains 1 fibronectin type III domain.		
CC	-1- SIMILARITY: Contains 49 immunoglobulin-like C2-type domains.		
CC	-1- SIMILARITY: Contains 1 PH domain.		
CC	-1- SIMILARITY: Contains 5 RCD5 domain.		
CC	-1- SIMILARITY: Contains 1 SH3 domain.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch</a> ).		
CC	-----		
DR	EMBL; U33058; AAB0542.1; -		
DR	EMBL; AF003131; AAB54132.2; -		
DR	PDB: 1PHO; 20-DEC-00.		
DR	Wormpep; C09D1.1; CE30426.		
DR	InterPro; IPR003961; FM_III.		
DR	InterPro; IPR007110; Ig-like.		
DR	InterPro; IPR003598; Ig-C2.		
DR	InterPro; IPR003006; Ig_MHC.		
DR	InterPro; IPR001849; PH.		
DR	InterPro; IPR007850; RCD5.		
DR	InterPro; IPR000219; RhGEF.		
DR	InterPro; IPR001452; SH3.		
DR	Pfam; PF00041; fn3; 1.		
DR	Pfam; PF00047; Ig; 47.		
DR	Pfam; PF00169; PH; 1.		
DR	Pfam; PF05177; RCD5; 5.		
DR	Pfam; PF00621; RhGEF; 1.		
DR	Pfam; PF00018; SH3; 1.		
DR	SMART; SM00408; IGC2; 23.		
DR	SMART; SM00325; RhGEF; 1.		
DR	SMART; SM00326; SH3; 1.		
DR	PROSITE; PS50010; DH_2; 1.		
DR	PROSITE; PS50835; IG-LIKE; 49.		
DR	PROSITE; PS50003; PH_DOMAIN; 1.		
DR	PROSITE; PS50002; SH3; 1.		
KW	Muscle protein; Immunoglobulin domain; Repeat; SH3 domain;		
KW	3D-structure.		
FT	DOMAIN 63 127 SH3.		
FT	DOMAIN 152 330 DH.		
FT	DOMAIN 342 498 PH.		
FT	DOMAIN 547 633 IG-LIKE C2-TYPE 1.		
FT	DOMAIN 648 736 IG-LIKE C2-TYPE 2.		
FT	DOMAIN 748 838 IG-LIKE C2-TYPE 3.		
FT	DOMAIN 946 1033 IG-LIKE C2-TYPE 4.		
FT	DOMAIN 1044 1132 IG-LIKE C2-TYPE 5.		
FT	DOMAIN 1140 1227 IG-LIKE C2-TYPE 6.		
FT	DOMAIN 1272 1315 THR-RICH.		
FT	DOMAIN 1375 1475 RCD5 1.		





Best Local Similarity 39.3%; Pred. No. 5.3;  
Matches 11; Conservative 6; Mismatches 6; Indels 5; Gaps 1;

QY 1 ORLKVAOEELSN-----VLPPPRKMG 23  
:::|:|:|:|:|:|:|:|:|  
Db 117 EKVKRAEELAEERYLDALPPPRDAG 144

Search completed: October 1, 2003, 09:57:55  
Job time : 9.23529 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 41.4706 Seconds  
(without alignments)  
155.564 Million cell updates/sec

Title: US-09-171-432a-47  
Perfect score: 129  
Sequence: 1 QRLKYAOEELSNEVLP...PPRKMKGLF 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues  
Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	94	12 Q9ENT9	Q9ent9 hepatitis a
2	129	100.0	94	12 Q9ENT3	Q9ent3 hepatitis a
3	129	100.0	94	12 Q9ENV9	Q9env9 hepatitis a
4	129	100.0	94	12 Q9ENV5	Q9env5 hepatitis a
5	129	100.0	94	12 Q9ENV2	Q9env2 hepatitis a
6	129	100.0	94	12 Q9ENT1	Q9ent1 hepatitis a
7	129	100.0	94	12 Q9ENT8	Q9ent8 hepatitis a
8	129	100.0	94	12 Q9ENV2	Q9env2 hepatitis a
9	129	100.0	94	12 Q9ENV2	Q9env2 hepatitis a
10	129	100.0	94	12 Q9ENV5	Q9env5 hepatitis a
11	129	100.0	94	12 Q9ENV1	Q9env1 hepatitis a
12	129	100.0	94	12 Q9ENV7	Q9env7 hepatitis a
13	129	100.0	94	12 Q9ENV7	Q9env7 hepatitis a
14	129	100.0	94	12 Q9ENV8	Q9env8 hepatitis a
15	129	100.0	94	12 Q9ENV6	Q9env6 hepatitis a
16	129	100.0	94	12 Q9ENV0	Q9env0 hepatitis a

17	129	100.0	94	12 Q9ENT6	Q9ent6 hepatitis a
18	129	100.0	94	12 Q9ENV4	Q9env4 hepatitis a
19	129	100.0	94	12 Q9ENV1	Q9env1 hepatitis a
20	129	100.0	94	12 Q9ENV5	Q9env5 hepatitis a
21	129	100.0	94	12 Q9ENV7	Q9env7 hepatitis a
22	129	100.0	94	12 Q9ENV4	Q9env4 hepatitis a
23	129	100.0	115	12 Q9DWR4	Q9dwr4 hepatitis a
24	129	100.0	116	12 Q92941	Q92941 hepatitis a
25	129	100.0	116	12 Q9W7X7	Q9w7x7 hepatitis a
26	129	100.0	116	12 Q9W7S7	Q9w7s7 hepatitis a
27	129	100.0	116	12 Q71977	Q71977 hepatitis a
28	129	100.0	116	12 Q8B8K6	Q8b8k6 hepatitis a
29	129	100.0	116	12 Q8B8K5	Q8b8k5 hepatitis a
30	129	100.0	116	12 Q8B8K4	Q8b8k4 hepatitis a
31	129	100.0	116	12 Q8B8K3	Q8b8k3 hepatitis a
32	129	100.0	126	12 Q9BMT9	Q9bmt9 hepatitis a
33	129	100.0	132	12 Q8V4L9	Q8v4l9 hepatitis a
34	129	100.0	132	12 Q8V4L6	Q8v4l6 hepatitis a
35	129	100.0	132	12 Q8V4J9	Q8v4j9 hepatitis a
36	129	100.0	132	12 Q8V4J5	Q8v4j5 hepatitis a
37	129	100.0	132	12 Q8V4M5	Q8v4m5 hepatitis a
38	129	100.0	132	12 Q8V4M7	Q8v4m7 hepatitis a
39	129	100.0	132	12 Q8V4N2	Q8v4n2 hepatitis a
40	129	100.0	132	12 Q8V4M4	Q8v4m4 hepatitis a
41	129	100.0	132	12 Q8V4J7	Q8v4j7 hepatitis a
42	129	100.0	132	12 Q8V4L5	Q8v4l5 hepatitis a
43	129	100.0	132	12 Q8V4K3	Q8v4k3 hepatitis a
44	129	100.0	132	12 Q8V4I0	Q8v4i0 hepatitis a
45	129	100.0	132	12 Q8V4M2	Q8v4m2 hepatitis a

## ALIGNMENTS

RESULT 1  
ID Q9ENT9 PRELIMINARY; PRT; 94 AA.  
AC Q9ENT9;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Nagasaki 32;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
RT infection: Prolonged HAV viremia and mild liver injury.\*;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB038298; BAB1836.1; -  
FT NON\_TER 1  
FT NON\_TER 94  
SQ SEQUENCE 94 AA; 10807 MW; F194CE91BB8C4AFD CRC64;  
Query Match 100.0%; Score 129; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 4.4e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QRLKYAOEELSNEVLP...PPRKMKGLF 25  
DB 61 QRLKYAOEELSNEVLP...PPRKMKGLF 85  
RESULT 2  
ID Q9ENV3 PRELIMINARY; PRT; 94 AA.  
AC Q9ENV3;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
DE 01-MAR-2001 (TReMBLrel. 16, last annotation update)  
OS Polypotein (Fragment).  
OC Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Nagasaki 27;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
RT infection: Prolonged HAV viremia and mild liver injury."  
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AB038294; BAB11832.1; -.  
FT NON\_TER 1 1  
FT SEQUENCE 94 94  
SQ SEQUENCE 94 AA; 10892 MW; 9AF8BE91BB8C53CC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 4.4e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPrKKKGLF 25  
DB 61 ORLKYAOEELSNEVLPPrKKKGLF 85

## RESULT 3

Q9ENV9 PRELIMINARY; PRT; 94 AA.  
AC 09ENV9;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Nagasaki 03;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
RT infection: Prolonged HAV viremia and mild liver injury."  
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AB038278; BAB11816.1; -.  
FT NON\_TER 1 1  
FT SEQUENCE 94 94  
SQ SEQUENCE 94 AA; 10862 MW; 9AF9EFD4A8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 4.4e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPrKKKGLF 25  
DB 61 ORLKYAOEELSNEVLPPrKKKGLF 85

## RESULT 4

Q9ENV5 PRELIMINARY; PRT; 94 AA.  
AC 09ENV5;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Nagasaki 08;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
RT infection: Prolonged HAV viremia and mild liver injury."  
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AB038282; BAB11820.1; -.  
FT NON\_TER 1 1  
FT SEQUENCE 94 94  
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 4.4e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPrKKKGLF 25  
DB 61 ORLKYAOEELSNEVLPPrKKKGLF 85

## RESULT 5

Q9ENV2 PRELIMINARY; PRT; 94 AA.  
AC 09ENV2;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Nagasaki 28;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
RT infection: Prolonged HAV viremia and mild liver injury."  
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AB038295; BAB11833.1; -.  
FT NON\_TER 1 1  
FT SEQUENCE 94 94  
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 4.4e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPrKKKGLF 25  
DB 61 ORLKYAOEELSNEVLPPrKKKGLF 85

## RESULT 6

Q9ENV1 PRELIMINARY; PRT; 94 AA.  
AC 09ENV1;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
DE Polypotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Kantou 49;



Query Match 100.0%; Score 129; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 4,4e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPrKKKGLF 25  
DB 61 ORLKYAOEELSNEVLPPrKKKGLF 85

## RESULT 11

Q9ENV1 PRELIMINARY; PRT; 94 AA.  
AC Q9ENV1; (Created)  
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
NCBI\_TaxID=12092;

SEQUENCE FROM N.A.  
RC STRAIN-Nagasaki 18;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
infection: Prolonged HAV viremia and mild liver injury.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB038286; BAB11824.1; -.  
FT NON\_TER 1  
FT 94  
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 4,4e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPrKKKGLF 25  
DB 61 ORLKYAOEELSNEVLPPrKKKGLF 85

## RESULT 12

Q9ENV9 PRELIMINARY; PRT; 94 AA.  
AC Q9ENV9; (Created)  
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
NCBI\_TaxID=12092;

SEQUENCE FROM N.A.  
RC STRAIN-Nagasaki 21;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
infection: Prolonged HAV viremia and mild liver injury.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB038288; BAB11826.1; -.  
FT NON\_TER 1  
FT 94  
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 4,4e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPrKKKGLF 25

DB 61 ORLKYAOEELSNEVLPPrKKKGLF 85

## RESULT 13

Q9ENV7 PRELIMINARY; PRT; 94 AA.  
AC Q9ENV7; (Created)  
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
NCBI\_TaxID=12092;

SEQUENCE FROM N.A.  
RC STRAIN-Nagasaki 06;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
infection: Prolonged HAV viremia and mild liver injury.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB038280; BAB11818.1; -.  
FT NON\_TER 1  
FT 94  
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 4,4e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPrKKKGLF 25  
DB 61 ORLKYAOEELSNEVLPPrKKKGLF 85

## RESULT 14

Q9ENV8 PRELIMINARY; PRT; 94 AA.  
AC Q9ENV8; (Created)  
DT 01-MAR-2001 (TReMBLrel. 16, last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
NCBI\_TaxID=12092;

SEQUENCE FROM N.A.  
RC STRAIN-Nagasaki 04;  
RA Ida S., Tachikawa N., Nakajima A., Daikoku M., Yano M., Kikuchi Y.,  
RA Yasuoka A., Oka S.;  
RT "Influence of HIV-1 infection on acute hepatitis A virus (HAV)  
infection: Prolonged HAV viremia and mild liver injury.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB038279; BAB11817.1; -.  
FT NON\_TER 1  
FT 94  
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91BB8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;  
Best Local Similarity 100.0%; Pred. No. 4,4e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPrKKKGLF 25  
DB 61 ORLKYAOEELSNEVLPPrKKKGLF 85

## RESULT 15

```

OGENV6
ID OGENV6 PRELIMINARY; PRT; 94 AA.
AC OGENV6;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE Polypeptide (Fragment).
OS Hepatitis A virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID-12092;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Nagasaki 07;
RA Ida S.;
RT *Influence of HIV-1 infection on acute hepatitis A virus (HAV)
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB038281; BAB1819.1; -.
NR NON_TER 1 1
SQ SEQUENCE 94 AA; 10876 MW; 9AF8BE91B8C4BAC CRC64;

Query Match 100.0%; Score 129; DB 12; Length 94;
Best Local Similarity 100.0%; Pred. No. 4.4e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ORLKTAQEELSNNEVLPPPRKMKGLF 25
DB 61 ORLKTAQEELSNNEVLPPPRKMKGLF 85

```

Search completed: October 1, 2003, 10:02:41  
 Job time : 41.4706 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 ; Search time 15.8824 seconds  
(without alignments)  
66,600 Million cell updates/sec

Title: US-09-171-432A-47  
Perfect score: 129  
Sequence: 1 QRLKYAOEELSNEVLPPPRKMKGLF 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues  
Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCRTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	1091	6	5516630-2
2	129	100.0	2227	3	US-08-475-886-2
3	129	100.0	2227	3	US-08-475-886-6
4	129	100.0	2227	3	US-08-397-232-2
5	129	100.0	2227	3	US-08-397-232-4
6	129	100.0	2227	3	US-09-171-387-2
7	129	100.0	2227	4	US-09-653-499-2
8	129	100.0	2227	4	US-09-653-499-6
9	124	96.1	2227	4	US-08-475-886-4
10	124	96.1	2227	4	US-09-653-499-4
11	117	90.7	839	4	US-08-087-016-2
12	49	38.0	686	4	US-09-328-352-4303
13	47	36.4	180	4	US-09-252-991A-18339
14	46.5	36.0	23	1	PCT-US91-03368-4
15	46.5	36.0	23	1	PCT-US91-03368-4
16	46.5	36.0	139	1	US-08-278-729A-8
17	46.5	36.0	139	1	US-08-155-343A-8
18	46.5	36.0	139	1	US-08-406-672-8
19	46.5	36.0	139	1	US-08-643-563A-8
20	46.5	36.0	139	1	US-08-643-563A-8
21	46.5	36.0	139	1	US-08-462-623-8
22	46.5	36.0	139	1	US-08-451-953A-8
23	46.5	36.0	139	2	US-08-445-468A-8
24	46.5	36.0	139	2	US-08-461-397A-8
25	46.5	36.0	139	2	US-08-912-088-8
26	46.5	36.0	139	3	US-08-278-730A-8
27	46.5	36.0	139	3	US-08-445-467-8

28	46.5	36.0	139	3	US-08-480-515A-8	Sequence 8, Appl1
29	46.5	36.0	139	3	US-08-414-033A-8	Sequence 8, Appl1
30	46.5	36.0	139	3	US-08-271-556A-6	Sequence 6, Appl1
31	46.5	36.0	139	3	US-08-440-894A-8	Sequence 8, Appl1
32	46.5	36.0	139	4	US-09-170-936-8	Sequence 8, Appl1
33	46.5	36.0	139	4	US-08-461-113-8	Sequence 8, Appl1
34	46.5	36.0	139	4	US-08-456-033-8	Sequence 8, Appl1
35	46.5	36.0	139	4	US-08-643-321-7	Sequence 7, Appl1
36	46.5	36.0	139	4	US-09-464-206-8	Sequence 8, Appl1
37	46.5	36.0	139	4	US-08-404-113A-8	Sequence 8, Appl1
38	46.5	36.0	139	5	PCT-US92-01968-8	Sequence 8, Appl1
39	46.5	36.0	139	5	PCT-US93-07150-8	Sequence 8, Appl1
40	46.5	36.0	139	5	PCT-US93-07231-8	Sequence 8, Appl1
41	46.5	36.0	139	5	PCT-US93-08742-8	Sequence 8, Appl1
42	46.5	36.0	139	5	PCT-US93-08808-8	Sequence 8, Appl1
43	46.5	36.0	139	5	PCT-US93-08885-8	Sequence 8, Appl1
44	46.5	36.0	397	1	US-07-841-646-27	Sequence 27, Appl1
45	46.5	36.0	397	5	PCT-US91-07635-4	Sequence 4, Appl1

ALIGNMENTS

```

RESULT 1
5516630-2
; Patent No. 5516630
; APPLICANT: TICEHURST, JOHN R.; BALTIMORE, DAVID; PEINSTONE,
; STEPHEN M.; PURCELL, ROBERT H.; RACANIELLO, VINCENT R.;
; BAROUDY, BAHIGE M.
; TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/788,262
; FILING DATE: 06-NOV-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 256,135
; FILING DATE: 06-OCT-1988
; APPLICATION NUMBER: 654,942
; FILING DATE: 27-SEP-1984
; APPLICATION NUMBER: 537,911
; FILING DATE: 30-SEP-1983
; SEQ ID NO:2:
; LENGTH: 1091
5516630-2

Query Match      100.0%; Score 129; DB 6; Length 1091;
Best Local Similarity 100.0%; Pred. No. 6.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 QRLKYAOEELSNEVLPPPRKMKGLF 25
DB      1047 QRLKYAOEELSNEVLPPPRKMKGLF 1071

RESULT 2
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE D
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
    
```

```

; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: MILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
; US-08-475-886-2

Query Match          100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAOEELSNEVLPPIPRKMKGLF 25
Db 810 QRLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 3
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated (4380) HAV, strain HM-175
; US-08-475-886-6

Query Match          100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAOEELSNEVLPPIPRKMKGLF 25
Db 810 QRLKYAOEELSNEVLPPIPRKMKGLF 834

ULT 4
; Sequence 2, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; EARLIER FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: MILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
; US-08-397-232-2
```

```

Query Match          100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAOEELSNEVLPPIPRKMKGLF 25
Db 810 QRLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 5
; Sequence 4, Application US/08397232A
; Patent No. 6180110
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US1
; CURRENT APPLICATION NUMBER: US/08/397,232A
; EARLIER FILING DATE: 1995-04-17
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: PCT/US93/08610
; EARLIER FILING DATE: 1993-09-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PRF
; ORGANISM: Attenuated (4380) HAV, strain HM-175
; US-08-397-232-4

Query Match          100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QRLKYAOEELSNEVLPPIPRKMKGLF 25
Db 810 QRLKYAOEELSNEVLPPIPRKMKGLF 834

RESULT 6
; Sequence 2, Application US/09171387
; Patent No. 6280734
; GENERAL INFORMATION:
; APPLICANT: RAYCHAUDHURI, GOPA;
; APPLICANT: EMERSON, SUZANNE, U.;
; APPLICANT: PURCELL, ROBERT, H.
; TITLE OF INVENTION: HAVING A CHIMERIC 2C PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/171,387
; FILING DATE: 24-Mar-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/06506
; FILING DATE: 18-Apr-1997
```

```
APPLICATION NUMBER: US60/015,642
FILING DATE: 19-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: William S. Feller
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-42290S1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 2
US-09-171-387-2

Query Match      100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ORLKTAQDELSEVLP...PPRKMGFLF 25
DB      810 ORLKTAQDELSEVLP...PPRKMGFLF 834

RESULT 7
US-09-653-499-2
Sequence 2, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

Query Match      100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ORLKTAQDELSEVLP...PPRKMGFLF 25
DB      810 ORLKTAQDELSEVLP...PPRKMGFLF 834

RESULT 8
US-09-653-499-6
Sequence 6, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
```

```
FILE REFERENCE: 202642620S2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

Query Match      100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.5e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 ORLKTAQDELSEVLP...PPRKMGFLF 25
DB      810 ORLKTAQDELSEVLP...PPRKMGFLF 834

RESULT 9
US-08-475-886-4
Sequence 4, Application US/08475886A
Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S2
CURRENT APPLICATION NUMBER: US/08/475,886A
CURRENT FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-08-475-886-4

Query Match      96.1%; Score 124; DB 3; Length 2227;
Best Local Similarity 96.0%; Pred. No. 8.9e-11;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      1 ORLKTAQDELSEVLP...PPRKMGFLF 25
DB      810 ORLKTAQDELSEVLP...PPRKMGFLF 834

RESULT 10
US-09-653-499-4
Sequence 4, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 202642620S2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31.
```

PRIOR APPLICATION NUMBER: 08/475,886  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 4  
LENGTH: 2227  
TYPE: PR  
ORGANISM: Attenuated HAV (pass 35), strain HM-175  
US-09-653-499-4

Query Match 96.1%; Score 124; DB 4; Length 2227;  
Best Local Similarity 96.0%; Pred. No. 8,9e-11;  
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPEPRKMKGLF 25  
|||||  
810 ORLKYAOEELSNEVLPPEPRKMKGLF 834

RESULT 11  
US-08-087-016-2  
Sequence 2, Application US/08087016  
Patent No. 5430135

GENERAL INFORMATION:  
APPLICANT: NAINAN, OMANA V.  
APPLICANT: MARGOLIS, HAROLD S.  
APPLICANT: ROBERTSON, BETTY H.  
APPLICANT: BRINTON, MARGO H.  
APPLICANT: EBERT, JAMES W.  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINE  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
STREET: 1615 L Street N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/087,016  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/678,828  
FILING DATE: 03-APR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: SCOTT, WATSON T.  
REGISTRATION NUMBER: 26,581  
REFERENCE/DOCKET NUMBER: WTS/5683/83834/SRL  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 861-3000  
TELEFAX: (202) 822-0944  
TELEX: 6714627 CUSH  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 839 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-087-016-2

Query Match 90.7%; Score 117; DB 1; Length 839;  
Best Local Similarity 84.0%; Pred. No. 3,8e-10;  
Matches 21; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPEPRKMKGLF 25  
|||||  
Db 809 ORFYAREELSNEVLPPEPRKMKGLF 833

RESULT 12  
US-09-328-352-4303  
Sequence 4303, Application US/09328352  
Patent No. 6562958  
GENERAL INFORMATION:  
APPLICANT: Gary L. Bretton et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTE  
FILE REFERENCE: GTC99-03PA  
CURRENT APPLICATION NUMBER: US/09/328,352  
CURRENT FILING DATE: 1999-06-04  
NUMBER OF SEQ ID NOS: 8252  
SEQ ID NO 4303  
LENGTH: 686  
TYPE: PR  
ORGANISM: Acinetobacter baumannii  
US-09-328-352-4303

Query Match 38.0%; Score 49; DB 4; Length 686;  
Best Local Similarity 41.4%; Pred. No. 14;  
Matches 12; Conservative 5; Mismatches 8; Indels 4; Gaps 1;

QY 1 ORLKYAOEELSNEVLPPEPRKMKGLF 25  
|||||  
Db 343 EKLKQOQFNLSQELKPYFPAPKYIGLF 371

RESULT 13  
US-09-252-991A-18339  
Sequence 18339, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
FILE REFERENCE: 107196,136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 18339  
LENGTH: 180  
TYPE: PR  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18339

Query Match 36.4%; Score 47; DB 4; Length 180;  
Best Local Similarity 43.5%; Pred. No. 6.7;  
Matches 10; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 ORLKYAOEELSNEVLPPEPRKMKG 23  
|||||  
Db 29 RLRLYAAORLQRLVLPQRRVFP 51

RESULT 14  
US-07-800-364B-4  
Sequence 4, Application US/07800364B  
Patent No. 5688678  
GENERAL INFORMATION:  
APPLICANT: Hewick, Rodney M.  
APPLICANT: Wang, Jack H.  
APPLICANT: Wozney, John M.  
APPLICANT: Celeste, Anthony J.

TITLE OF INVENTION: Bone and Cartilage Inductive Proteins  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Legal Affairs, Genetics Institute, Inc.  
 STREET: 87 Cambridgepark Drive  
 CITY: Cambridge  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02140  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/800,364B  
 FILING DATE: 26-NOV-1991  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kapinos, Ellen J.  
 REGISTRATION NUMBER: 32,245  
 REFERENCE/DOCKET NUMBER: GI 5182A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-876-1170  
 TELEFAX: 617-876-5851  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 23 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Bos taurus  
 TISSUE TYPE: Bone  
 US-07-800-364B-4

Query Match 36.0%; Score 46.5; DB 1; Length 23;  
 Best Local Similarity 60.0%; Pred. No. 0.84;  
 Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 11 SNEVLPPPRKMKGLF 25  
 :||| ||| | :|||  
 Db 1 TNE-LPPNKLPGIF 14

SUITE 15  
 US91-03388-4  
 Sequence 4, Application PC/TUS9103388  
 GENERAL INFORMATION:  
 APPLICANT: Hewick, Rodney M.  
 TITLE OF INVENTION: Bone and Cartilage Inductive Proteins  
 NUMBER OF SEQUENCES: 13  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Legal Affairs, Genetics Institute, Inc.  
 STREET: 87 Cambridgepark Drive  
 CITY: Cambridge  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02140  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US91/03388  
 FILING DATE: 19910515  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kapinos, Ellen J.

REGISTRATION NUMBER: 32,245  
 REFERENCE/DOCKET NUMBER: GI5182X-PCT  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-876-1170  
 TELEFAX: 617-876-5851  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 23 amino acids  
 TYPE: AMINO ACID  
 STRANDEDNESS: single  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Bos taurus  
 TISSUE TYPE: Bone  
 PCT-US91-03388-4

Query Match 36.0%; Score 46.5; DB 5; Length 23;  
 Best Local Similarity 60.0%; Pred. No. 0.84;  
 Matches 9; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 11 SNEVLPPPRKMKGLF 25  
 :||| ||| | :|||  
 Db 1 TNE-LPPNKLPGIF 14

Search completed: October 1, 2003, 10:06:34  
 Job time : 16.8824 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 ; Search time 30.1471 Seconds  
(without alignments)  
131.201 Million cell updates/sec

Title: US-09-171-432a-47

Perfect score: 129

Sequence: 1 ORLKYAOEELSNEVLPPPRKMKGLF 25

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCV\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/CTUS\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/CTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	352	15	US-10-272-459-45
2	129	100.0	980	15	US-10-272-459-41
3	129	100.0	2227	10	US-09-929-955-12
4	129	100.0	2227	14	US-10-104-966-12
5	129	100.0	2227	14	US-10-135-988-2
6	129	100.0	2227	14	US-10-135-988-6
7	129	100.0	836	15	US-10-272-459-40
8	124	96.1	2227	14	US-10-135-988-4
9	46.5	36.0	139	8	US-08-260-675-8
10	46.5	36.0	139	15	US-10-050-050-8
11	46.5	36.0	399	8	US-08-957-425-27
12	46.5	36.0	399	8	US-08-260-675-23
13	46.5	36.0	399	12	US-10-350-747-2
14	46.5	36.0	399	15	US-10-122-026-8
15	46.5	36.0	399	15	US-10-050-050-23

16	46	35.7	263	9	US-09-864-761-37656	Sequence 37656, A
17	45.5	35.3	1596	11	US-09-909-5678-47	Sequence 47, Appl
18	45	34.9	223	10	US-09-895-513A-84	Sequence 84, Appl
19	45	34.9	445	15	US-10-177-293-296	Sequence 296, App
20	43.5	33.7	222	15	US-10-156-761-11163	Sequence 11163, A
21	43.5	33.7	1114	12	US-09-840-743-14	Sequence 14, Appl
22	43	33.3	311	15	US-10-156-761-10491	Sequence 10491, A
23	43	33.3	442	15	US-10-156-761-13833	Sequence 13833, A
24	43	33.3	592	12	US-10-100-294A-36	Sequence 36, Appl
25	43	33.3	869	9	US-09-815-242-5230	Sequence 5230, Ap
26	43	33.3	882	9	US-09-815-242-12526	Sequence 12526, A
27	43	33.3	1095	15	US-10-128-714-3039	Sequence 3039, Ap
28	43	33.3	1209	15	US-10-128-714-3144	Sequence 3144, Ap
29	43	33.3	1255	15	US-10-128-714-8144	Sequence 8144, Ap
30	43	33.3	1277	15	US-10-128-714-8039	Sequence 8039, Ap
31	43	33.3	1370	11	US-09-842-758-111	Sequence 111, App
32	43	33.3	1570	11	US-09-842-758-35	Sequence 35, Appl
33	43	33.3	1648	10	US-09-515-806-4	Sequence 4, Appl1
34	43	33.3	1648	11	US-09-842-758-37	Sequence 37, Appl
35	43	33.3	1648	11	US-09-842-758-39	Sequence 39, Appl
36	43	33.3	2696	12	US-10-309-933-4	Sequence 4, Appl1
37	42	32.6	88	9	US-09-864-761-35516	Sequence 35516, A
38	42	32.6	88	9	US-09-864-761-43148	Sequence 43148, A
39	42	32.6	93	10	US-09-925-300-1101	Sequence 1101, Ap
40	42	32.6	146	10	US-09-923-246-56	Sequence 56, Appl
41	42	32.6	146	10	US-09-825-561A-47	Sequence 47, Appl
42	42	32.6	146	15	US-10-295-723-56	Sequence 56, Appl
43	42	32.6	192	14	US-10-001-857-119	Sequence 119, App
44	42	32.6	223	15	US-10-205-823-101	Sequence 101, Appl
45	42	32.6	311	8	US-08-818-581B-6	Sequence 6, Appl1

#### ALIGNMENTS

RESULT 1  
US-10-272-459-45  
; Sequence 45, Application US/10272459  
; Publication No. US20030124517A1  
; GENERAL INFORMATION:  
; APPLICANT: PICHOANTES, Sergio  
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT  
; TITLE OF INVENTION: PROTEINS AND USES THEREOF  
; FILE REFERENCE: P17955.002 / 2301-17955  
; CURRENT APPLICATION NUMBER: US/10/272,459  
; CURRENT FILING DATE: 2002-10-15  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein  
; OTHER INFORMATION: of 38.8 kDa  
US-10-272-459-45

Query Match 100.0% Score 129; DB 15; Length 352;  
Best Local Similarity 100.0%; Pred. No. 1.1e-10;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 ORLKYAOEELSNEVLPPPRKMKGLF 25  
DB 182 ORLKYAOEELSNEVLPPPRKMKGLF 206  
RESULT 2  
US-10-272-459-41  
; Sequence 41, Application US/10272459  
; Publication No. US20030124517A1  
; GENERAL INFORMATION:  
; APPLICANT: PICHOANTES, Sergio  
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

```
; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI/955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272.459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41
```

```
Query Match          100.0%; Score 129; DB 15; Length 980;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
1 ORLKYAOEELSNVLPPEPRKMGLE 25
|||||
810 ORLKYAOEELSNVLPPEPRKMGLE 834
```

```
RESULT 3
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-09-929-955-12
```

```
Query Match          100.0%; Score 129; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ORLKYAOEELSNVLPPEPRKMGLE 25
|||||
DB 810 ORLKYAOEELSNVLPPEPRKMGLE 834
```

```
RESULT 4
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
```

```
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12
```

```
Query Match          100.0%; Score 129; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ORLKYAOEELSNVLPPEPRKMGLE 25
|||||
DB 810 ORLKYAOEELSNVLPPEPRKMGLE 834
```

```
RESULT 5
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2
```

```
Query Match          100.0%; Score 129; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ORLKYAOEELSNVLPPEPRKMGLE 25
|||||
DB 810 ORLKYAOEELSNVLPPEPRKMGLE 834
```

```
RESULT 6
US-10-135-988-6
; Sequence 6, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
```



PRIOR FILING DATE: 1995-03-10  
 NUMBER OF SEQ ID NOS: 6  
 SOFTWARE: Patent In Ver. 2.1  
 SEQ ID NO 6  
 LENGTH: 2227  
 TYPE: PR  
 ORGANISM: Attenuated (4380) HAV, strain HM-175  
 US-10-135-988-6

Query Match  
 Best Local Similarity 100.0%; Score 129; DB 14;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-10;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORLKYAQEELSNEVLPPPRKMKGLF 25  
 DB 810 ORLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 7  
 US-10-272-459-40  
 Sequence 40, Application US/10272459  
 Publication No. US20030124517A1

GENERAL INFORMATION:  
 APPLICANT: PICHANTES, Sergio  
 TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT  
 TITLE OF INVENTION: PROTEINS AND USES THEREOF  
 FILE REFERENCE: P17955.002 / 2301-17955  
 CURRENT APPLICATION NUMBER: US/10/272,459  
 CURRENT FILING DATE: 2002-10-15  
 NUMBER OF SEQ ID NOS: 50  
 SOFTWARE: Patent In Ver. 2.0  
 SEQ ID NO 40  
 LENGTH: 836  
 TYPE: PR  
 ORGANISM: Artificial Sequence  
 FEATURE: OTHER INFORMATION: Description of Artificial Sequence: recombinant protein  
 OTHER INFORMATION: of 94 kDa  
 US-10-272-459-40

Query Match  
 Best Local Similarity 96.9%; Score 125; DB 15;  
 Best Local Similarity 96.0%; Pred. No. 1e-09;  
 Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORLKYAQEELSNEVLPPPRKMKGLF 25  
 DB 810 ORLKYAQEELSNEVLPPPRKMKGLF 834

US-10-135-988-4  
 Sequence 4, Application US/10135988  
 Publication No. US20020176869A1  
 GENERAL INFORMATION:  
 APPLICANT: FUNKHOUSER, ANN W  
 APPLICANT: EMERSON, SUZANNE U  
 APPLICANT: PURCELL, ROBERT H  
 APPLICANT: D'HONDT, ERIC  
 TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
 FILE REFERENCE: 20264262053  
 CURRENT APPLICATION NUMBER: US/10/135,988  
 CURRENT FILING DATE: 2002-04-29  
 PRIOR APPLICATION NUMBER: 07/947,338  
 PRIOR FILING DATE: 1992-09-18  
 PRIOR APPLICATION NUMBER: 08/397,232  
 PRIOR FILING DATE: 1995-03-10  
 NUMBER OF SEQ ID NOS: 6  
 SOFTWARE: Patent In Ver. 2.1  
 SEQ ID NO 4  
 LENGTH: 2227  
 TYPE: PR  
 ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
 US-10-135-988-4

Query Match  
 Best Local Similarity 96.1%; Score 124; DB 14;  
 Best Local Similarity 96.0%; Pred. No. 4.2e-09;  
 Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ORLKYAQEELSNEVLPPPRKMKGLF 25  
 DB 810 ORLKYAQEELSNEVLPPPRKMKGLF 834

RESULT 9  
 US-08-260-675-8  
 Sequence 8, Application US/08260675  
 Publication No. US20030104993A1

GENERAL INFORMATION:  
 APPLICANT: RUBGER, DAVID C  
 APPLICANT: KUBERASAMPATH, THANGAVEL  
 APPLICANT: OPPERMAN, HERMANN  
 APPLICANT: OZKAYNAK, ENGIN  
 APPLICANT: PANG, ROY HL  
 APPLICANT: COHEN, CHARLES M  
 TITLE OF INVENTION: MORPHOGEN-INDUCED NERVE REGENERATION AND  
 TITLE OF INVENTION: REPAIR  
 NUMBER OF SEQUENCES: 33  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: TESTA HURWITZ & THIBEAULT  
 STREET: 55 STATE STREET  
 CITY: BOSTON  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02140

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/260,675  
 FILING DATE:  
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/126,100  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/07/922,813  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/667,274  
 FILING DATE: 11-MAR-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/752,764  
 FILING DATE: 30-AUG-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: PITCHER ESQ, EDMUND R  
 REGISTRATION NUMBER: 27,829  
 REFERENCE/DOCKET NUMBER: CRP-070  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617/248-7000  
 TELEFAX: 617/248-7100

INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 139 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 ORIGINAL SOURCE:  
 ORGANISM: MORIDAE  
 TISSUE TYPE: EMBRYO  
 FEATURE:  
 NAME/KEY: Protein  
 LOCATION: 1..139  
 OTHER INFORMATION: /label= MOP2-MATURE

US-08-260-675-8

Query Match 36.0%; Score 46.5; DB 8; Length 139;  
Best Local Similarity 47.8%; Pred. No. 35;  
Matches 11; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 3 LKYOELSNVLPPrKMGlf 25  
DB 5 LKRQPKTNE-LPHPNKLPgIf 26

RESULT 10

US-10-050-050-8  
; Sequence 0, Application US/10050050  
; Publication No. US20030125230A1

; GENERAL INFORMATION:

; APPLICANT: COHEN, CHARLES M.

; KUBERASAMPATH, THANGAVEL

; RUEGER, DAVID C.

; OPPERMANN, HERMANN

; PANG, ROY H.L.

; OZKAYNAK, ENGIN

; SMART, JOHN E.

; TITLE OF INVENTION: MORPHOGEN TREATMENT FOR LIMITING

; PROLIFERATION OF EPITHELIAL CELLS.

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES

; STREET: 45 SOUTH STREET

; CITY: HOPKINTON

; STATE: MA

; COUNTRY: USA

; ZIP: 01748

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/050.050

; FILING DATE: 15-Jan-2002

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/461.113

; FILING DATE: &lt;Unknown&gt;

; APPLICATION NUMBER: US 08/445.882

; FILING DATE: 22-MAY-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: FENTON ESQ., GILLIAN M.

; REGISTRATION NUMBER: 36,508

; REFERENCE/DOCKET NUMBER: CRP-074DV

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (508) 435-9001

; TELEFAX: (508) 435-6951

; INFORMATION FOR SEQ ID NO: /note- "MOP-2 (MATURE FORM)"

; SEQUENCE CHARACTERISTICS:

; LENGTH: 139 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; NAME/KEY: Protein

; LOCATION: 1..139

; SEQUENCE DESCRIPTION: SEQ ID NO: 8:

DB 5 LKRQPKTNE-LPHPNKLPgIf 26

RESULT 11  
US-08-957-425-27  
; Sequence 27, Application US/08957425  
; Publication No. US20030069401A1

; GENERAL INFORMATION:

; APPLICANT: OPPERMANN, HERMANN

; KUBERASAMPATH, THANGAVEL

; RUEGER, DAVID C.

; PANG, ROY H.L.

; TITLE OF INVENTION: OSTEOGENIC DEVICES

; NUMBER OF SEQUENCES: 33

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: TESTA, HURWITZ &amp; THIBEAULT

; STREET: 53 STATE STREET

; CITY: BOSTON

; STATE: MASSACHUSETTS

; COUNTRY: U.S.A.

; ZIP: 02109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/957.425

; FILING DATE: 24-Oct-1997

; CLASSIFICATION: &lt;Unknown&gt;

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/447.570

; FILING DATE: 21-FEB-1992

; APPLICATION NUMBER: US 810.560

; FILING DATE: 20-DEC-1991

; APPLICATION NUMBER: US 827.052

; FILING DATE: 28-JAN-1992

; APPLICATION NUMBER: US 660.162

; FILING DATE: 22-FEB-1991

; APPLICATION NUMBER: US 621.988

; FILING DATE: 04-DEC-1990

; APPLICATION NUMBER: US 621.849

; FILING DATE: 04-DEC-1990

; APPLICATION NUMBER: US 616.374

; FILING DATE: 21-NOV-1990

; APPLICATION NUMBER: US 600.024

; FILING DATE: 18-OCT-1990

; APPLICATION NUMBER: US 599.543

; FILING DATE: 18-OCT-1990

; APPLICATION NUMBER: US 579.865

; FILING DATE: 07-SEP-1990

; APPLICATION NUMBER: US 569.920

; FILING DATE: 20-AUG-1990

; APPLICATION NUMBER: US 483.913

; FILING DATE: 22-FEB-1990

; APPLICATION NUMBER: US 422.613

; FILING DATE: 17-OCT-1989

; APPLICATION NUMBER: US 315.342

; FILING DATE: 23-FEB-1988

; APPLICATION NUMBER: US 232.630

; FILING DATE: 15-AUG-1988

; APPLICATION NUMBER: US 179.460

; FILING DATE: 08-APR-1988

; ATTORNEY/AGENT INFORMATION:

; NAME: PITCHER, EDMUND R.

; REGISTRATION NUMBER: 27,829

; REFERENCE/DOCKET NUMBER: CRP-001CP6

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/248-7000

; TELEFAX: 617/248-7100

; INFORMATION FOR SEQ ID NO: 27:

QY 3 LKYOELSNVLPPrKMGlf 25

Query Match 36.0%; Score 46.5; DB 15; Length 139;

Best Local Similarity 47.8%; Pred. No. 35;

Matches 11; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

SEQUENCE CHARACTERISTICS:  
LENGTH: 399 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
US-08-957-425-27

Query Match 36.0%; Score 46.5; DB 8; Length 399;  
Best Local Similarity 47.8%; Pred. No. 1.1e+02;  
Matches 11; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 3 LKTAQELSNVLPKPKKGLF 25  
DB 265 LKRRQPKKTNE-LPHPNKLPKGF 286

## RESULT 12

US-08-260-675-23  
Sequence 23, Application US/08260675  
Publication No. US2003010493A1

## GENERAL INFORMATION:

APPLICANT: RUEGER, DAVID C  
APPLICANT: KUBERASAMPATH, THANGAVEL  
APPLICANT: OPPERMAN, HERMANN  
APPLICANT: OKAYNAK, ENGIN  
APPLICANT: PANG, ROY HL  
APPLICANT: COHEN, CHARLES M  
TITLE OF INVENTION: MORPHOGEN-INDUCED NERVE REGENERATION AND  
TITLE OF INVENTION: REPAIR  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESS: TESTA HURWITZ & THIBEAULT  
STREET: 55 STATE STREET  
CITY: BOSTON  
STATE: MA  
COUNTRY: USA  
ZIP: 02140

## COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/260,675  
FILING DATE:  
CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/126,100  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/922,813  
FILING DATE:  
APPLICATION DATA:  
APPLICATION NUMBER: US/07/667,274  
FILING DATE: 11-MAR-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/752,764  
FILING DATE: 30-AUG-1991

## ATTORNEY/AGENT INFORMATION:

NAME: FITCHER ESQ, EDMUND R  
REGISTRATION NUMBER: 27,829  
REFERENCE/DOCKET NUMBER: CRP-070  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/248-7000  
TELEFAX: 617/248-7100

## INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:  
LENGTH: 399 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-260-675-23

Query Match 36.0%; Score 46.5; DB 8; Length 399;  
Best Local Similarity 47.8%; Pred. No. 1.1e+02;  
Matches 11; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 3 LKTAQELSNVLPKPKKGLF 25  
DB 265 LKRRQPKKTNE-LPHPNKLPKGF 286

## RESULT 13

US-10-350-747-2  
Sequence 2, Application US/10350747  
Publication No. US20030153072A1

## GENERAL INFORMATION:

APPLICANT: Hogan, Bridgid L.M.  
TITLE OF INVENTION: Compositions and Methods of Making  
Embryonic Stem Cells  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Panitch Schwarze Jacobs & Nadel, P.C.  
STREET: 1601 Market Street, 36th Floor  
CITY: Philadelphia  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19103-2398

## COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/350,747  
FILING DATE: 24-Jan-2003  
CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/808,346  
FILING DATE: 28-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Leary Ph.D., Kathryn  
REGISTRATION NUMBER: 36,317  
REFERENCE/DOCKET NUMBER: 9823-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-567-2020  
TELEFAX: 215-567-2991  
TELEX: 831-494

## INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 399 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-10-350-747-2

Query Match 36.0%; Score 46.5; DB 12; Length 399;  
Best Local Similarity 47.8%; Pred. No. 1.1e+02;  
Matches 11; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

QY 3 LKTAQELSNVLPKPKKGLF 25  
DB 265 LKRRQPKKTNE-LPHPNKLPKGF 286

## RESULT 14

US-10-122-026-8  
Sequence 8, Application US/10122026  
Publication No. US20030105004A1

## GENERAL INFORMATION:

APPLICANT: JONES, WILLIAM K  
TUCKER, RONALD F  
RUEGER, DAVID C



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:30 : Search time 54.4118 Seconds  
(without alignments)  
72.928 Million cell updates/sec

Title: US-09-171-432a-48  
Perfect score: 129  
Sequence: 1 WLNPKKINLADRMGLSGVOEIKQ 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues 1107863  
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database :  
1: A.Geneseq.10Jun03.\*  
2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*  
24: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*  
25: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	25	18	AAW42969
2	129	100.0	25	22	ABG31728
3	129	100.0	2227	7	ABG31729
4	129	100.0	2227	11	ABG31730
5	129	100.0	2227	18	ABG31731
6	129	100.0	2227	21	ABG31732
7	129	100.0	2227	21	ABG31733
8	129	100.0	2227	21	ABG31734
9	129	100.0	2227	23	ABG31735

10	129	100.0	2227	23	ABG31728
11	129	100.0	2227	23	ABG31729
12	129	100.0	2227	23	ABG31730
13	129	100.0	2227	24	ABG31731
14	129	100.0	2227	24	ABG31732
15	129	100.0	2227	24	ABG31733
16	129	100.0	2227	24	ABG31734
17	129	100.0	2227	24	ABG31735
18	129	100.0	2227	24	ABG31736
19	129	100.0	2227	24	ABG31737
20	129	100.0	2227	24	ABG31738
21	129	100.0	2227	24	ABG31739
22	129	100.0	2227	24	ABG31740
23	129	100.0	2227	24	ABG31741
24	129	100.0	2227	24	ABG31742
25	129	100.0	2227	24	ABG31743
26	129	100.0	2227	24	ABG31744
27	129	100.0	2227	24	ABG31745
28	129	100.0	2227	24	ABG31746
29	129	100.0	2227	24	ABG31747
30	129	100.0	2227	24	ABG31748
31	129	100.0	2227	24	ABG31749
32	129	100.0	2227	24	ABG31750
33	129	100.0	2227	24	ABG31751
34	129	100.0	2227	24	ABG31752
35	129	100.0	2227	24	ABG31753
36	129	100.0	2227	24	ABG31754
37	129	100.0	2227	24	ABG31755
38	129	100.0	2227	24	ABG31756
39	129	100.0	2227	24	ABG31757
40	129	100.0	2227	24	ABG31758
41	129	100.0	2227	24	ABG31759
42	129	100.0	2227	24	ABG31760
43	129	100.0	2227	24	ABG31761
44	129	100.0	2227	24	ABG31762
45	129	100.0	2227	24	ABG31763

## ALIGNMENTS

RESULT 1	AAW42969	standard; peptide; 25 AA.
ID	AAW42969	
XX	AAW42969	
AC	AAW42969	
XX	AAW42969	
DT	28-APR-1998	(first entry)
XX	28-APR-1998	
DE	Immunogenic Hepatitis A virus peptide YK-1757.	
XX	Immunogenic Hepatitis A virus peptide; p2a protein;	
KW	Immunogenic response; antibody.	
XX	Immunogenic response; antibody.	
OS	Synthetic.	
OS	Hepatitis A virus.	
XX	Hepatitis A virus.	
PN	WO9740147-A1.	
XX	WO9740147-A1.	
PD	30-OCT-1997.	
XX	30-OCT-1997.	
PF	18-APR-1997;	97WO-US06891.
XX	18-APR-1997;	
PR	19-APR-1996;	96US-0015644.
XX	19-APR-1996;	
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICES.	
XX	(USSH ) US DEPT HEALTH & HUMAN SERVICES.	
PI	Fields HA, Khudyakov YE;	
XX	Fields HA, Khudyakov YE;	
DR	WPI; 1997-535831/49.	
XX	WPI; 1997-535831/49.	
PT	Immunogenic Hepatitis A virus (HAV) peptide(s) - used to induce an	
PT	Immune response to HAV in a mammal or to detect the presence of	

Hepatitis A virus  
Attenuated Hepatitis A virus  
Wild type human be  
Attenuated (pass3  
Immunogenic Hepat  
Synthetic HAV p2a  
S. lividans protea  
Amino acid sequenc  
Human kinase (PKIN  
Human steroid and  
Drosophila melanog  
Saccharomyces cere  
Human FLEXHT-45 pr  
Human T2R41 amino  
Human novel secret  
Human novel polype  
Human protein sequ  
Human FLEXHT-16 pr  
S. epidermidis ope  
Human novel secret  
Human novel polype  
Staphylococcus epi  
Novel human diagno  
Human polypeptide  
Novel human diagno  
Recombinant Group  
Arabidopsis thaliana  
Arabidopsis thaliana  
Arabidopsis thaliana

PT antibodies against HAV in a mammal  
XX  
PS Claim 18; Page 112; 140pp; English.  
XX  
CC The present immunogenic peptide corresponds to an immunogenic  
CC epitope of the Hepatitis A virus (HAV). The peptide is substantially  
CC similar to a portion of the amino acid sequence of the P2A protein of HAV  
CC corresponding to amino acids 792-980. Compositions containing the  
CC peptide can be used to induce an immune response to HAV in a mammal.  
CC The peptide can also be used to detect the presence of antibodies  
CC against HAV in mammalian serum. The peptide can also be used to make an  
CC antibody against HAV by administering the peptide to a mammal.  
XX  
SQ Sequence 25 AA:  
Query Match 100.0%; Score 129; DB 18; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1.4e-14;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 WLNPRTINLADRMGLSGVDEIKRQ 25  
1 WLNPRTINLADRMGLSGVDEIKRQ 25  
RESULT 2  
AAB69448  
ID AAB69448 standard; Peptide: 25 AA.  
XX  
AC AAB69448;  
XX  
DT 20-APR-2001 (first entry)  
XX  
DE Synthetic HAV P2A peptide, SEQ ID NO: 48.  
XX  
KW Hepatitis A virus; HAV; Immunogen; Immunostimulant; Virucide; Vaccine;  
KW antigen; major structural capsid polypeptide; HAV antibody detection.  
XX  
OS Hepatitis A virus.  
OS Synthetic.  
XX  
PN WO200105824-A2.  
XX  
PD 25-JAN-2001.  
XX  
PF 14-JUL-2000; 2000WO-US19267.  
XX  
PR 15-JUL-1999; 99US-0144412.  
XX  
(USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
Fields HA, Khudiyakov YE;  
XX  
DR WPI: 2001-112681/12.  
XX  
PT Synthetic peptides used as antigen sources for enzyme immunoassays  
PT detecting anti-hepatitis A virus and as vaccines -  
XX  
PS Claim 13; Page 99; 130pp; English.  
XX  
CC The present sequence is one of a number of synthetic peptides which are  
CC immunoreactive with hepatitis A virus (HAV) antibodies. The peptides  
CC comprise antigenic epitopes of the major structural capsid polypeptides  
CC or non-structural polypeptides of HAV with one or more glutamine  
CC molecules at the carboxy end of the peptide. The peptides are used to  
CC detect the presence of antibodies against HAV in mammalian serum, to  
CC detect the presence of HAV in a human or animal through the binding of  
CC the peptide to an antibody, to detect acute phase infection by detecting  
CC IgM antibodies in mammalian serum and detecting convalescence in a  
CC mammal. The peptides are used to detect or quantify HAV antibodies in  
CC samples in clinical or research-based assays using immunoblotting,  
CC fluorescence in situ hybridisation analysis, gel-mobility shift assays,  
CC tracking of radioactive or bioluminescent markers, chromatography or  
CC electrophoresis. The peptides are used to induce an immune response to

CC HAV when administered to a human or animal. Glutamine at the carboxy  
CC end of the peptides enhances the IgM antibody reactivity.  
XX  
SQ Sequence 25 AA:  
Query Match 100.0%; Score 129; DB 22; Length 25;  
Best Local Similarity 100.0%; Pred. No. 1.4e-14;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 WLNPRTINLADRMGLSGVDEIKRQ 25  
1 WLNPRTINLADRMGLSGVDEIKRQ 25  
Db  
RESULT 3  
AAP60066  
ID AAP60066 standard; Protein: 2227 AA.  
XX  
AC AAP60066;  
XX  
DT 25-MAR-2003 (updated)  
DT 26-JUN-1991 (first entry)  
XX  
DE Sequence of viral I434 polypeptide encoded by the complete  
DE nucleotide sequence of the HAV genome.  
XX  
KW Diagnosis; vaccine; passive immunotherapy.  
XX  
OS Hepatitis A virus.  
XX  
FH Key  
FH Region Location/Qualifiers  
FT 1..245  
FT /label= P1.1A  
FT Region 246..491  
FT /label= 1B  
FT Region 492..836  
FT /label= 1C  
FT Region 837..980  
FT /label= P2.2A  
FT Region 981..1076  
FT /label= 2B  
FT Region 1077..1422  
FT /label= 2C  
FT Region 1423..1484  
FT /label= P3.3A  
FT Region 1485..1507  
FT /label= 3B  
FT Region 1508..1678  
FT /label= 3C  
FT Region 1679..2227  
FT /label= 3D  
XX  
PN EP19480-A.  
XX  
XX 29-OCT-1986.  
XX  
PD 03-APR-1986; 86EP-0302465.  
XX  
PF 03-APR-1985; 85US-0719329.  
XX  
PR (CHIR ) CHIRON CORP.  
XX  
PI Dina D, Potter SJ, Vannest GA, Caput D;  
XX WPI: 1986-286213/44.  
XX N-PSDB; AAN60080.  
XX  
DR Hepatitis A virus nucleotide sequence and polypeptide - and use  
DR in prodn. of vaccines and diagnostic probes  
XX  
PS Claim 5; Fig 1; 18pp; English.  
XX  
CC AAN60080 and oligonucleotide fragments are useful in detection of

hepatitis A virus; transformed hosts may be used for expression of  
CC polypeptides and fragments useful in vaccines without risk of  
CC infection by the virus or in prodn. of particles which are capable  
CC of inducing immunocompetent B cells for passive immunotherapy. Pref.  
CC epitope is derived from AAs 445-657 or 792-848 of the HAV  
CC polypeptide sequence (AAP60066).  
CC (Updated on 25-MAR-2003 to correct PA field.)

SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 7; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 3.1e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIRKEQ 25  
|||||  
DB 956 WLNPKKINLADRMGLSGVOEIRKEQ 980

ULT 4  
05697  
AAR05697 standard; protein; 2227 AA.

AC AAR05697;  
XX  
XX 25-MAR-2003 (updated)  
DT 15-AUG-1990 (first entry)

DE Attenuated hepatitis A virus.

KW Hepatitis A virus; vaccine; attenuated.

OS Hepatitis A virus; strain HM-175.

FH Key Location/Qualifiers

FT 1..23 /label-VP4 = 1A

FT 24..245 /label-VP2 = 1B

FT 246..491 /label-VP3 = 1C

FT 492..791 /label-VP1 = 1D

FT 792..980 /label-VP1 = 1D

FT 981..1087 /label-2A

FT 1088..1422 /label-2B

FT 1423..1496 /label-2C

FT 1497..1519 /label-3A

FT 1520..1738 /label-3B = VPg

FT 1739..2227 /label-3C

FT 1739..2227 /label-3D

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

FT Region

DR N-PSDB; AAO03512.  
XX  
XX Vaccine against hepatitis A virus infection - comprises novel  
PT attenuated hepatitis A virus strain.  
XX  
XX Claim 1; Fig 1; 18pp; English.  
XX

The attenuated HAV is useful for inducing protective immunity against  
CC HAV. This strain (Pass 35) differs from the wild type HAV HM-175 by  
CC several nucleotide changes distributed throughout the genome, is  
CC attenuated for chimpanzees, elicits serum neutralizing antibodies, and is  
CC suitable for use as an HAV vaccine. It is noted that not all the changes  
CC are necessary for attenuation and use as a vaccine.  
CC (Updated on 25-MAR-2003 to correct PA field.)  
CC (Updated on 25-MAR-2003 to correct PI field.)

SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 11; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 3.1e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIRKEQ 25  
|||||  
DB 956 WLNPKKINLADRMGLSGVOEIRKEQ 980

RESULT 5  
AAM34074  
ID AAM34074 standard; protein; 2227 AA.

AC AAM34074;

DT 27-APR-1998 (first entry)

DE Hepatitis A virus HM-175 protein sequence.

KW HAV; attenuation; 2C protein; 2C chimeric hepatitis A virus;

XX infection; vaccine.

OS Hepatitis A virus HM-175.

FH Key Location/Qualifiers

FT 1..23 /label-VP4

FT 24..245 /label-VP2

FT 246..491 /label-VP3

FT 492..791 /label-VP1

FT 792..980 /label-2A

FT 981..1087 /label-2A

FT 1088..1422 /label-2B

FT 1423..1496 /label-2C

FT 1497..1519 /label-3A

FT 1520..1738 /label-3B

FT 1739..2227 /label-3C

FT 1739..2227 /label-3D

FT Protein

FT Protein

FT Protein

FT Protein

FT Protein

WO9740166-A2.  
XX  
XX 30-OCT-1997.  
XX  
XX 18-APR-1997; 97WO-US06506.  
XX  
XX 19-APR-1996; 96US-0015642.

XX (USSH ) US SEC DEPT HEALTH.  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Emerson SU, Purcell RH, Raychaudhuri G;  
XX  
DR WPI: 1997-535850/49.  
DR N-PSDB: AAT93023.  
PT Human attenuated HAV genome containing simian HAV 2C gene - useful  
XX as vaccines against HAV infection  
PS Disclosure; Fig 13A-D; 66pp; English.  
XX  
CC This protein sequence is encoded by the human hepatitis A virus  
CC (HAV) HM-175 wild-type genome (see AAT93023). Attenuated strain  
CC HAV/7 is obtained by passage of HM-175 in African Green Monkey  
CC kidney cells. A claimed DNA construct (I) comprises a genome of  
CC HAV, where the genome is a human attenuated HAV genome in which a  
CC region of the 2C gene has been replaced by a corresponding region  
CC from a 2C gene of a simian AGM-27 HAV genome (see AAT93024). The  
CC region of the 2C gene from AGM-27 contained in the construct  
CC preferably encodes amino acids 120-328 of the 2C protein, amino  
CC acids 1-121 or amino acids 1-328. Also claimed are: (1) an RNA  
CC transcript of (1); (2) a cell transfected with (1) or the RNA  
CC transcript of (1); (3) a HAV genome as above; (4) antibodies to the  
CC HAV of (3); and (5) a host cell containing the HAV of (3). (1) or  
CC its RNA transcript, can be used as a vaccine for preventing HAV in  
CC a mammal. (1) or the RNA transcript can also be used to stimulate  
CC the production of protective antibodies in the mammal.  
XX  
SQ Sequence 2227 AA:  
Query Match 100.0%; Score 129; DB 18; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 3,1e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 WLNPKRINLADRMIGLSGVOEIKQ 25  
DB 956 WLNPKRINLADRMIGLSGVOEIKQ 980  
RESULT 6  
AAB18607  
ID AAB18607 standard; Protein; 2227 AA.  
AC AAB18607;  
XX  
XX 15-JAN-2001 (first entry)  
XX Amino acid sequence of wild type Hepatitis A virus strain HM-175.  
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection.  
XX Hepatitis A virus.  
XX US6113912-A.  
XX  
XX 05-SEP-2000.  
XX  
XX 07-JUN-1995; 95US-0475886.  
XX  
XX 18-SEP-1992; 92US-0947338.  
XX 17-SEP-1993; 93WO-US08610.  
XX 10-MAR-1995; 95US-0397232.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;  
XX WPI: 2000-586464/55.  
XX N-PSDB: AAA75476.  
XX

PT Novel live hepatitis A virus adapted to growth in human fibroblast cell  
PT line useful as vaccine for protecting humans against hepatitis A virus  
PT infection, has modified genome compared to wild type  
XX  
XX Disclosure; Fig 6A-K; 72pp; English.  
XX  
XX The present sequence is derived from a wild type hepatitis A virus  
XX (HAV) strain HM-174. The sequence is modified to produce HAV which  
XX are adapted to growth in the human fibroblast-like cell line MRC-5.  
XX The HAV is able to propagate in MRC-5 cells and retain appropriate  
XX attenuation. It is useful as a live vaccine for prophylaxis of  
XX hepatitis A in humans and other primates.  
XX  
SQ Sequence 2227 AA:  
Query Match 100.0%; Score 129; DB 21; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 3,1e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 WLNPKRINLADRMIGLSGVOEIKQ 25  
DB 956 WLNPKRINLADRMIGLSGVOEIKQ 980  
RESULT 7  
AAB18608  
ID AAB18608 standard; Protein; 2227 AA.  
AC AAB18608;  
XX  
XX 15-JAN-2001 (first entry)  
XX  
XX Amino acid sequence of passage 35 Hepatitis A virus called P-35.  
XX  
XX HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;  
XX P-35 virus.  
XX  
XX Hepatitis A virus.  
XX  
XX US6113912-A.  
XX  
XX 05-SEP-2000.  
XX  
XX 07-JUN-1995; 95US-0475886.  
XX  
XX 18-SEP-1992; 92US-0947338.  
XX 17-SEP-1993; 93WO-US08610.  
XX 10-MAR-1995; 95US-0397232.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX D'Hondt E, Purcell RH, Emerson SU, Funkhouser AW;  
XX WPI: 2000-586464/55.  
XX N-PSDB: AAA75477.  
XX  
XX Novel live hepatitis A virus adapted to growth in human fibroblast cell  
XX line useful as vaccine for protecting humans against hepatitis A virus  
XX infection, has modified genome compared to wild type  
XX  
XX Disclosure; Columns 67-78; 72pp; English.  
XX  
XX The present sequence is derived from passage 35 of a wild type  
XX hepatitis A virus (HAV) strain HM-174. The resulting virus is  
XX designated P-35 virus. The sequence is modified to produce HAV which  
XX are adapted to growth in the human fibroblast-like cell line MRC-5.  
XX The HAV is able to propagate in MRC-5 cells and retain appropriate  
XX attenuation. It is useful as a live vaccine for prophylaxis of  
XX hepatitis A in humans and other primates.  
XX  
SQ Sequence 2227 AA:  
Query Match 100.0%; Score 129; DB 21; Length 2227;



Best Local Similarity 100.0%; Pred. No. 3.1e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

# RESULT 8

AAB18609  
ID AAB18609 standard; Protein; 2227 AA.

AC AAB18609;

DT 15-JAN-2001 (first entry)

XX Amino acid sequence of live attenuated Hepatitis A virus 4380.

DE HAV; strain HM-174; MRC-5 cell; live vaccine; hepatitis A infection;  
HM HAV 4380.

XX Hepatitis A virus.

XX US6113912-A.

PD 05-SEP-2000.

PF 07-JUN-1995; 95US-0475886.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93WO-US08610.

PR 10-MAR-1995; 95US-0397232.

PA (USSR ) US DEPT HEALTH & HUMAN SERVICES.

PI D'Hondt E, Purcell RH, Emerson SU, Funkhouser AM;

DR WPI: 2000-586464/55.

DR N-PSDB: AAA75478.

XX Novel live hepatitis A virus adapted to growth in human fibroblast cell  
XX line useful as vaccine for protecting humans against hepatitis A virus  
XX infection, has modified genome compared to wild type

XX Disclosure; Columns 93-104; 72pp; English.

CC The present sequence is derived from a live attenuated hepatitis A  
CC virus (HAV) of the invention, designated HAV 4380. The sequence is  
CC produced by modifying wild type HAV strain HM-174. The HAV of the  
CC invention are adapted to growth in the human fibroblast-like cell  
CC line MRC-5. The HAV is able to propagate in MRC-5 cells and retain  
CC appropriate attenuation. It is useful as a live vaccine for prophylaxis  
CC of hepatitis A in humans and other primates.

XX Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 21; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 3.1e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

# RESULT 9

ABG31727  
ID ABG31727 standard; Protein; 2227 AA.

AC ABG31727;

DT 29-NOV-2002 (first entry)

DE Wild-type Hepatitis A virus (HAV) strain HM-175 polypeptide.

XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide.

XX Hepatitis A virus strain HM-175.

XX US6423318-B1.

XX 23-JUL-2002.

PF 31-ANG-2000; 2000US-0653499.

PR 07-JUN-1995; 95US-0475886.

PR 17-SEP-1993; 93US-0397232.

PR 17-SEP-1993; 93WO-US08610.

XX (USSR ) US DEPT HEALTH & HUMAN SERVICES.

PA (SMK ) SMITHKLINE BEECHAM BIOLOGICALS.

PI Funkhouser AM, Emerson SU, Purcell RH, D'Hondt E;

DR WPI: 2002-680946/73.

DR N-PSDB: ABS2787.

XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth  
XX in MRC-5 cells, useful for preparing a vaccine against HAV infection

XX Disclosure; Fig 6; 71pp; English.

CC The invention relates to a polynucleotide which encodes a hepatitis A  
CC virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell  
CC line). The polynucleotide is useful for preparing a vaccine against  
CC hepatitis A virus infection. This sequence represents a hepatitis A virus  
CC strain HM-175 polypeptide.

XX Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 23; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 3.1e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

# RESULT 10

ABG31728  
ID ABG31728 standard; Protein; 2227 AA.

AC ABG31728;

DT 29-NOV-2002 (first entry)

DE Hepatitis A virus mutant strain HM-175/7 (PHAV/7) polypeptide.

XX Hepatitis A virus; HAV; HM-175; HM-175/7; MRC-5 cell; human fibroblast;

KW virucide; mutant; PHAV/7; muten.

XX Hepatitis A virus strain HM-175.

OS Synthetic.

XX Key

FT MISC-difference 963 Location/Qualifiers

FT MISC-difference 764 /label= Wild-type Lys substituted by Arg

FT MISC-difference 821 /note= "Wild-type Glu substituted by Val"

FT MISC-difference 1052 /note= "Wild-type Asn substituted by Ser"

FT MISC-difference 1062 /note= "Wild-type Ala substituted by Val"

FT MISC-difference /note= "Wild-type Gly substituted by Ala"

```
FT Misc-difference 1118 /note= "Wild-type Lys substituted by Met"
FT Misc-difference 1151 /note= "Wild-type Glu substituted by Lys"
FT Misc-difference 1163 /note= "Wild-type Phe substituted by Ser"
FT Misc-difference 1277 /note= "Wild-type Val substituted by Ile"
FT Misc-difference 1500 /note= "Wild-type His substituted by Tyr"
FT Misc-difference 1805 /note= "Wild-type Asp substituted by Asn"
FT Misc-difference 1930 /note= "Wild-type Ser substituted by Thr"
FT
XX
XX US6423318-B1.
XX
XX 23-JUL-2002.
XX
XX 31-AUG-2000; 2000US-0653499.
XX
XX 07-JUN-1995; 95US-0475886.
XX 17-SEP-1993; 93US-0397232.
XX 17-SEP-1993; 93WO-US08610.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
XX WPI: 2002-680946/73.
XX DR N-PSDB; ABS52788.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX in MRC-5 cells, useful for preparing a vaccine against HAV infection -
XX
XX Example 3; Column 67-78; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
XX virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX line). The polynucleotide is useful for preparing a vaccine against
XX hepatitis A virus infection. This sequence represents a hepatitis A virus
XX mutant strain HM-175/7 (pHAV/7) polypeptide.
XX
XX Sequence 2227 AA:
XX
XX Query Match 100.0%; Score 129; DB 23; Length 2227;
XX Best Local Similarity 100.0%; Pred. No. 3.1e-12;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 WLNPRTKINLADRMGLSGVQEIKEQ 25
XX I|||||
XX 956 WLNPRTKINLADRMGLSGVQEIKEQ 980
XX
XX Db
XX
XX RESULT 11
XX ABG31729
XX ID ABG31729 standard; Protein; 2227 AA.
XX
XX ABG31729;
XX
XX 29-NOV-2002 (first entry)
XX
XX Attenuated Hepatitis A virus (HAV) 4380 polypeptide.
XX
XX Hepatitis A virus; HAV; HM-175; MRC-5 cell; human fibroblast; virucide;
XX HAV 4380.
XX
XX Hepatitis A virus strain HM-175.
XX
XX US6423318-B1.
XX
XX 23-JUL-2002.
XX
```

```
XX
XX 31-AUG-2000; 2000US-0653499.
XX
XX 07-JUN-1995; 95US-0475886.
XX 17-SEP-1993; 93US-0397232.
XX 17-SEP-1993; 93WO-US08610.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;
XX WPI: 2002-680946/73.
XX DR N-PSDB; ABS52789.
XX
XX New polynucleotide encoding a hepatitis A virus (HAV) adapted to growth
XX in MRC-5 cells, useful for preparing a vaccine against HAV infection -
XX
XX Disclosure; Column 93-104; 71pp; English.
XX
XX The invention relates to a polynucleotide which encodes a hepatitis A
XX virus (HAV) adapted to growth in MRC-5 cells (human fibroblast-like cell
XX line). The polynucleotide is useful for preparing a vaccine against
XX hepatitis A virus infection. This sequence represents an attenuated
XX hepatitis A virus 4830 polypeptide.
XX
XX Sequence 2227 AA:
XX
XX Query Match 100.0%; Score 129; DB 23; Length 2227;
XX Best Local Similarity 100.0%; Pred. No. 3.1e-12;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 WLNPRTKINLADRMGLSGVQEIKEQ 25
XX I|||||
XX 956 WLNPRTKINLADRMGLSGVQEIKEQ 980
XX
XX Db
XX
XX RESULT 12
XX AAE19899
XX ID AAE19899 standard; Protein; 2227 AA.
XX
XX AAE19899;
XX
XX 18-JUN-2002 (first entry)
XX
XX Hepatitis A virus (HAV) protein.
XX
XX Hepatitis A virus; HAV; infection; virucide; fungicide; antibacterial;
XX cytosstatic; immunostimulant; vaccine; ribavirin; immune response; cancer.
XX
XX Hepatitis A virus.
XX
XX WO200213855-A2.
XX
XX 21-FEB-2002.
XX
XX 15-AUG-2001; 2001WO-1B01808.
XX
XX 17-AUG-2000; 2000US-225767P.
XX 29-AUG-2000; 2000US-228175P.
XX 03-NOV-2000; 2000US-0705547.
XX
XX (TRIP-) TRIPER AB.
XX
XX Sallberg M, Hultgren C;
XX WPI: 2002-241837/29.
XX DR N-PSDB; AAD31766.
XX
XX Vaccine compositions for treating and preventing disease, preferably
XX hepatitis C virus infection, comprises ribavirin and antigen that has
XX epitope present in hepatitis C virus
```

PS Claim 11; Page 82-87; 120pp; English.  
 CC The invention relates to a composition comprising ribavirin and an  
 CC antigen preferably non structural 3 protein (NS3)/4A fragment of  
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV  
 CC sequence. The composition is useful for enhancing an immune response to  
 CC a hepatitis C antigen in humans, domestic, sport or pet species and as  
 CC vaccines for treating and preventing HCV infections. The composition is  
 CC also useful for treating viral, bacterial, fungal diseases and cancer.  
 CC The present sequence is hepatitis A virus (HAV) protein.  
 SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 23; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 3,1e-12;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25  
 ||||||||||||||||||||||||||||  
 956 WLNPKKINLADRMGLSGVQEIKEQ 980

## RESULT 13

ID ABU08639 standard; Protein; 2227 AA.

AC ABU08639;

DT 03-JUN-2003 (first entry)

DE Wild type human hepatitis A virus strain HM-175.

KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;

KM vaccine; MRC-5 cell; hepatitis infection.

OS Hepatitis A virus strain HM-175.

PN US2002176869-A1.

PD 28-NOV-2002.

PF 29-APR-2002; 2002US-0135988.

PR 07-JUN-1995; 95US-0475886.

PR 31-AUG-2000; 2000US-0653499.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93MO-US08610.

PR 17-APR-1995; 95US-0397232.

(FUNK/) FUNKHOUSER A W.

(EMER/) EMERSON S U.

(PURC/) PURCELL R H.

(DHON/) D'HONDT E.

PA FUNKHOUSER AM, Emerson SU, Purcell RH, D'Hondt E;

PI FUNKHOUSER AM, Emerson SU, Purcell RH, D'Hondt E;

PT WPI: 2003-352605/33.

DR N-PSDB; ABX93473.

XX Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 24; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 3,1e-12;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25  
 ||||||||||||||||||||||||||||  
 DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

## RESULT 14

ID ABU08640 standard; Protein; 2227 AA.

AC ABU08640;

DT 03-JUN-2003 (first entry)

DE Attenuated (pass35) hepatitis A virus strain HM-175.

KW Hepatitis A virus; HAV; virucide; hepatotropic; antiinflammatory;

KM vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.

OS Hepatitis A virus strain HM-175.

PN US2002176869-A1.

PD 28-NOV-2002.

PF 29-APR-2002; 2002US-0135988.

PR 07-JUN-1995; 95US-0475886.

PR 31-AUG-2000; 2000US-0653499.

PR 18-SEP-1992; 92US-0947338.

PR 17-SEP-1993; 93MO-US08610.

PR 17-APR-1995; 95US-0397232.

(FUNK/) FUNKHOUSER A W.

(EMER/) EMERSON S U.

(PURC/) PURCELL R H.

(DHON/) D'HONDT E.

PA FUNKHOUSER AM, Emerson SU, Purcell RH, D'Hondt E;

PI FUNKHOUSER AM, Emerson SU, Purcell RH, D'Hondt E;

PT WPI: 2003-352605/33.

DR N-PSDB; ABX93474.

XX Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 24; Length 2227;  
 Best Local Similarity 100.0%; Pred. No. 3,1e-12;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQEIKEQ 25  
 ||||||||||||||||||||||||||||  
 DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

## RESULT 15

ID ABU08641 standard; Protein; 2227 AA.

XX

AC AB008641;  
 XX  
 DT 03-JUN-2003 (first entry)  
 XX  
 DE Attenuated hepatitis A virus (4380) strain HM-175.  
 XX  
 KM Hepatitis A virus; HAV, virucide; hepatotropic; antiinflammatory;  
 XX vaccine; MRC-5 cell; hepatitis infection; ds; attenuated virus.  
 OS Hepatitis A virus strain HM-175.  
 XX  
 PN US2002176869-A1.  
 XX  
 PD 28-NOV-2002.  
 XX  
 PF 29-APR-2002; 2002US-0135988.  
 XX  
 PR 07-JUN-1995; 95US-0475886.  
 PR 31-AUG-2000; 2000US-0653499.  
 PR 18-SEP-1992; 92US-0947338.  
 PR 17-SEP-1993; 93WO-US08610.  
 PR 17-APR-1995; 95US-0397232.  
 XX  
 PA (FUNK/) FUNKHOUSER A W.  
 PA (EMER/) EMERSON S U.  
 PA (FURC/) FURCELL R H.  
 PA (DHON/) D'HONDT E.  
 PI Funkhouser AW, Emerson SU, Purcell RH, D'Hondt E;  
 XX  
 DR WPI: 2003-352605/33.  
 DR N-PSDB; ABX93475.  
 XX  
 PT New live hepatitis A virus, HAV4380, adapted to growth in MRC-5 cells,  
 PT useful in vaccines for protecting primates against hepatitis infection  
 PT and disease -  
 XX  
 PS Disclosure; Page 45-51; 70pp; English.  
 XX  
 CC The invention describes a live hepatitis A virus (HAV) adapted to growth  
 CC in MRC-5 cells. The HAV and compositions comprising the HAV are useful as  
 CC a vaccine for protecting primates against hepatitis infection and  
 CC disease. This is the amino acid sequence of an attenuated human  
 CC hepatitis A virus (4380) strain HM-175.  
 XX  
 SQ Sequence 2227 AA;

Query Match 100.0%; Score 129; DB 24; Length 2227;  
 Best Local Similarity 100.0%; Pred. NO. 3.1e-12;  
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKRQ 25  
 ||||||||||||||||||||  
 Db 956 WLNPKKINLADRMGLSGVOEIKRQ 980

Search completed: October 1, 2003, 09:56:47  
 Job time : 55.418 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:49:15 ; Search time 16.0294 Seconds  
(without alignments)  
149.988 Million cell updates/sec

Title: US-09-171-432a-48  
Perfect score: 129  
Sequence: 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues  
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	2227	1 GNNYHM	genome polyprotein
2	129	100.0	2227	1 GNNYHR	genome polyprotein
3	129	100.0	2227	1 GNNYMK	genome polyprotein
4	129	100.0	2227	1 GNNYHB	genome polyprotein
5	129	100.0	2230	1 GNNYSA	genome polyprotein
6	51.5	39.9	656	2 B82056	glutathione-regula
7	50	38.8	172	2 S33416	heat shock protein
8	49	38.0	327	1 RDVZAS	ribonucleoside-dip
9	48	37.2	158	2 S64321	hypothetical prote
10	47.5	36.8	482	2 S27608	cysteine proteinas
11	47.5	36.8	653	2 G82971	probable ferredoxi
12	47	36.4	145	2 B25199	hypothetical prote
13	47	36.4	381	2 T18876	hypothetical prote
14	46.5	36.0	146	2 B70142	ribosomal protein
15	46	35.7	342	2 H81317	probable lipopolys
16	46	35.7	342	2 H86336	hypothetical prote
17	46	35.7	518	2 S73432	MG096 homolog D09
18	46	35.7	632	2 S73431	MG288 homolog D09
19	46	35.7	671	2 AF0042	probable oxidoredu
20	46	35.7	877	2 S58824	probable membrane
21	45.5	35.3	201	2 H81274	amidotransferase C
22	45.5	35.3	1027	2 T43024	ceos protein - Bur
23	45	34.9	1042	2 H70203	iso-leucine-tRNA 11
24	45	34.9	1061	2 H90084	hypothetical prote
25	44	34.1	152	2 S07505	endoxyribonucle
26	44	34.1	165	2 S73194	hypothetical prote
27	44	34.1	278	2 B82388	probable maltose o
28	44	34.1	370	2 H70423	oxygen-independent
29	44	34.1	408	2 AC2295	succinyl-CoA synth

30	44	34.1	412	2 C96816	hypothetical prote
31	44	34.1	414	2 T06303	enoyl-CoA hydratase
32	44	34.1	512	2 T48462	cytochrome P450-11
33	44	34.1	646	2 S72609	GTP-binding membra
34	44	34.1	653	2 G70683	probable lepa - My
35	44	34.1	677	2 B82870	DNA topoisomerase
36	44	34.1	788	2 S67595	hypothetical prote
37	44	34.1	994	2 S18739	env protein - siml
38	44	34.1	1770	2 S56221	hypothetical prote
39	43.5	33.7	388	1 S72995	alanine racemase (
40	43.5	33.7	395	2 AD0380	multidrug efflux p
41	43.5	33.7	482	2 A38533	transcription acti
42	43	33.3	145	2 B24289	heat shock protein
43	43	33.3	148	2 G69125	ribosomal protein
44	43	33.3	180	2 AC0056	conserved hypothet
45	43	33.3	207	2 C43905	protein secretion

ALIGNMENTS

RESULT 1

GNNYHM

N: genome polyprotein - human hepatitis A virus (strain HM-175, wild type)

B: RNA-directed RNA polymerase (PC 2.7.7.48), protein 3D

C: Species: human hepatitis A virus

A: Note: host Homo sapiens (man)

C: Date: 31-Mar-1988 #sequence\_revision 31-Mar-1988 #text\_change 16-Jul-1999

C: Accession: A25981

R: Cohen, J.I.; Ticehurst, J.R.; Purcell, R.H.; Buckler-White, A.; Baroudy, B.M.

J. Virol. 61, 50-59, 1987

A: Title: Complete nucleotide sequence of wild-type hepatitis A virus: comparison with

A: Reference number: A25981; PMID:87061253; PMID:3023706

A: Accession: A25981

A: Molecule type: genomic RNA

A: Residues: 1-2227 <COH>

A: Cross-references: EMBL:M14707; NID:g329582; PIDN:AAA45465.1; PID:g329583

C: Superfamily: hepatitis A virus genome polyprotein

C: Keywords: coat protein; core protein; cysteine proteinase; hydrolase; nucleotidyltr

F:1-23/Product: coat protein 1A #status predicted <VP2>

F:24-245/Product: coat protein 1B #status predicted <VP2>

F:246-491/Product: coat protein 1C #status predicted <VP3>

F:492-791/Product: coat protein 1D #status predicted <VP1>

F:792-980/Product: core protein 2A #status predicted <C2A>

F:981-1087/Product: core protein 2B #status predicted <C2B>

F:1088-1422/Product: core protein 2C #status predicted <C2C>

F:1423-1496/Product: protein 3A #status predicted <C3A>

F:1497-1519/Product: protein 3B #status predicted <C3B>

F:1520-1738/Product: cysteine proteinase, protein 3C #status predicted <C3C>

F:1739-2227/Product: RNA-directed RNA polymerase, protein 3D #status predicted <C3D>

Query Match 100.0%; Score 129; DB 1; Length 2227;

Best Local Similarity 100.0%; Pred. No. 1.1e-10;

Matches .25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVQEIKEQ 25

DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 2

GNNYHR

N: genome polyprotein - human hepatitis A virus

B: RNA polymerase (PC 2.7.7.48), protein 3D

C: Species: human hepatitis A virus

A: Note: host Homo sapiens (man)

C: Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 16-Jul-1999

C: Accession: A03903

R: Najarian, R.; Caput, D.; Gee, W.; Potter, S.J.; Renard, A.; Merryweather, J.; van N

Proc. Natl. Acad. Sci. U.S.A. 82, 2627-2631, 1985

A: Title: Primary structure and gene organization of human hepatitis A virus.



F:28-249/Product: coat protein 1B #status predicted <C1B>  
F:250-495/Product: coat protein 1C #status predicted <C1C>  
F:96-795/Product: coat protein 1D #status predicted <C1D>  
F:796-984/Product: core protein 2A #status predicted <C2A>  
F:985-1091/Product: core protein 2B #status predicted <C2B>  
F:1092-1426/Product: core protein 2C #status predicted <C2C>  
F:1427-1498/Product: protein 3A #status predicted <P3A>  
F:1499-1521/Product: protein 3B #status predicted <P3B>  
F:1522-1741/Product: protein 3C #status predicted <P3C>  
F:1742-2230/Product: protein 3D #status predicted <P3D>

Query Match 100.0%; Score 129; DB 1; Length 2230;  
Best Local Similarity 100.0%; Pred. No. 1,1e-10;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLPKKNLADRMGLSGVOEIKQ 25  
Db 960 WLPKKNLADRMGLSGVOEIKQ 984

RESULT 6

glutathione-regulated potassium-efflux system protein KeftB VC2606 [Imported] - Vibrio cholerae  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: B82056  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.; Chardson, D.; Esmolaeva, M.D.; Vamathavan, J.; Baas, S.; Qian, H.; Dragol, I.; Sellers, F. I., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: AB2035; MID:20406833; PMID:10952301  
A:Accession: B82056  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-656 <HEI>  
A:Cross-references: GB:AE004327; GB:AE003852; NID:g9657185; PIDN:AAF95747.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC2606  
A:Map position: 1  
C:Superfamily: glutathione-regulated potassium efflux system protein keftC

Query Match 39.9%; Score 51.5; DB 2; Length 656;  
Best Local Similarity 54.5%; Pred. No. 13;  
Matches 12; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

2 LNPKKI-NLADRMGLSGVOEI 22  
133 LNPKKLMDLRGPIILGLGCAQV 154

RESULT 7

heat shock protein hsp20 - nematode (Nippostrongylus brasiliensis)  
C:Species: Nippostrongylus brasiliensis  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 13-Aug-1999  
C:Accession: S33416  
R:Tweedie, S.; Grigg, M.E.; Ingram, L.; Selkirk, M.E.  
A:Submitted to the EMBL Data Library, April 1993  
A:Description: The expression of a small heat shock homologue is developmentally regulated  
A:Reference number: S33416  
A:Accession: S33416  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-172 <TWB>  
A:Cross-references: EMBL:X71663; NID:g297865; PIDN:CAA50655.1; PID:g297866  
C:Superfamily: alpha-crystallin

Query Match 38.8%; Score 50; DB 2; Length 172;  
Best Local Similarity 43.5%; Pred. No. 5.2;  
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 3 NPKKINLADRMGLSGVOEIKQ 25  
Db 77 NELKVLDDRDILVEGMQVKTE 99

RESULT 8

ribonucleoside-diphosphate reductase (EC 1.17.4.1) small chain - African swine fever  
N:Alternate names: ribonucleotide reductase small chain  
C:Species: African swine fever virus, ASFV  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 06-Nov-1998  
C:Accession: B40568  
R:Boursnell, M.; Shaw, K.; Yanez, R.J.; Vinuela, E.; Dixon, L.  
A:Title: The sequences of the ribonucleotide reductase genes from African swine fever  
A:Reference number: A40568; MID:91335775; PMID:1871576  
A:Accession: B40568  
A:Molecule type: DNA  
A:Residues: 1-327 <BOU>  
A:Cross-references: GB:M64728  
C:Function:  
A:Description: catalyzes the reduction of ribonucleoside diphosphate to deoxyribonucleoside  
C:Superfamily: herpesvirus ribonucleoside-diphosphate reductase small chain  
C:Keywords: deoxyribonucleotide biosynthesis; early protein; iron; metalloprotein; ox  
F:70,101,104,164,196,201/Binding site: 2Fe-O cluster (Asp, Glu, His, Glu, His) #  
F:108/Active site: Tyr (stable tyrosyl radical) #status predicted

Query Match 38.0%; Score 49; DB 1; Length 327;  
Best Local Similarity 31.8%; Pred. No. 14;  
Matches 7; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

OY 1 WLPKKNLADRMGLSGVOEI 22  
Db 145 WMDPARNSLGERLYGFAAVEGI 166

RESULT 9

hypothetical protein YGR030C - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein G4068  
C:Species: Saccharomyces cerevisiae  
C:Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 19-Apr-2002  
C:Accession: S64321  
R:Rieger, M.; Mueller-Auer, S.; Bruckner, M.; Schaefer, M.  
A:Submitted to the Protein Sequence Database, May 1996  
A:Reference number: S64071  
A:Accession: S64321  
A:Molecule type: DNA  
A:Residues: 1-158 <RIE>  
A:Cross-references: EMBL:Z72815; NID:g1323007; PID:e243932; PID:g1323008; GSPDB:GN000  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:POF6; MRP5:YGR030C  
A:Cross-references: SGD:S0003262  
A:Map position: 7R

Query Match 37.2%; Score 48; DB 2; Length 158;  
Best Local Similarity 66.7%; Pred. No. 9.4;  
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 KRINLADRMGLSGV 19  
Db 64 KOINMADRLSLGLQV 78

RESULT 10

cysteine proteinase tpr - Porphyromonas gingivalis  
C:Species: Porphyromonas gingivalis  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 04-Mar-2000  
C:Accession: S27608  
R:Bourgeois, G.; Lapointe, H.; Ploquin, P.; Mayrand, D.  
A:Submitted to the EMBL Data Library, February 1992





A: Molecule type: DNA  
A: Residues: 1-342 <PAR>  
A: Cross-references: GB:AL139077; GB:AL111168; NID:g6968444; PIDN:CAB73388.1; PID:g696856  
A: Experimental source: serotype O2, strain NCTC 11168  
C: Genetics:  
A: Gene: waac: Cj1133  
C: Superfamily: ADP-heptose-LPS heptosyltransferase II

Query Match 35.7%; Score 46; DB 2; Length 342;  
Best Local Similarity 42.9%; Pred. No. 42;  
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVQE 21  
DB 220 WGNVKEVEFAKEVLNLSGIDE 240

Search completed: October 1, 2003, 10:04:48  
Job time : 18.0294 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:41:55 ; Search time 8.23529 Seconds

(without alignments)  
142.760 Million cell updates/sec

Title: US-09-171-432a-48

Perfect score: 129

Sequence: 1 WLNPKKINLADRMGLSGVQEIKEQ 25

Scoring table: BLOSUM62

Searched: 127863 seqs, 47026705 residues

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Database: SwissProt\_41:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	2226	1	P26580 hepatitis a
2	129	100.0	2226	1	P26581 hepatitis a
3	129	100.0	2226	1	P26582 hepatitis a
4	129	100.0	2227	1	P08617 hepatitis a
5	129	100.0	2227	1	P06441 hepatitis a
6	129	100.0	2227	1	P13901 hepatitis a
7	129	100.0	2230	1	P14553 simian hepa
8	50	38.8	172	1	Q07160 nipostrom
9	49	38.0	327	1	P26713 african swi
10	48	37.2	158	1	P53218 saccharomyc
11	48	37.2	334	1	P42492 african swi
12	48	37.2	622	1	O9rcf6 streptomyc
13	47.5	36.8	482	1	P25806 porphyromon
14	47	36.4	145	1	P06582 caenorhabd1
15	47	36.4	145	1	P42170 caenorhabd1
16	46.5	36.0	146	1	O51314 borrelia bu
17	46	35.7	518	1	P75066 mycoplasma
18	46	35.7	632	1	P75066 mycoplasma
19	45.5	35.3	201	1	O9pny4 campylobact
20	45	34.9	1042	1	O51773 borrelia bu
21	44	34.1	152	1	P20314 bacterioph
22	44	34.1	165	1	P51370 porphyra pu
23	44	34.1	646	1	P53330 mycobacteri
24	44	34.1	653	1	P71739 mycobacteri
25	44	34.1	985	1	P23073 simian fo
26	44	34.1	1770	1	P43365 saccharomyc
27	43.5	33.7	285	1	O8x876 buchnera ap
28	43.5	33.7	388	1	P38056 mycobacteri
29	43.5	33.7	441	1	O8rfx6 fuscobacteri
30	43.5	33.7	482	1	P29267 algaligenes
31	43	33.3	145	1	P34696 caenorhabd1
32	43	33.3	148	1	O26129 methanobact
33	43	33.3	180	1	O8zip1 yersinia pe

34	43	33.3	207	1	GSPJ_AERHY	P31739 aeromonas h
35	43	33.3	367	1	MX12_MOUSE	O08911 mus musculu
36	43	33.3	421	1	Y162_METJA	O57626 methanococc
37	43	33.3	495	1	NUSA_HAEIN	P43915 haemophilus
38	43	33.3	810	1	SYRA_SYNY3	P74396 synechocyst
39	42.5	32.9	284	1	RP32_BUCAL	O05385 buchnera ap
40	42.5	32.9	602	1	LEPA_BRUME	O8ydb8 bruceella me
41	42.5	32.9	687	1	ILVA_YEAST	P07342 saccharomyc
42	42.5	32.9	1472	1	ATC9_YEAST	O12697 saccharomyc
43	42	32.6	143	1	HS16_CAEL	P05581 caenorhabd1
44	42	32.6	151	1	SOPC_HALRO	P81926 halocynthia
45	42	32.6	185	1	TD52_MOUSE	O62393 mus musculu

## ALIGNMENTS

RESULT 1	ID	POLG_HPAV2	STANDARD:	PRT: 2226 AA.
AC	P26580;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].			
DE	Hepatitis A virus (strain 24a).			
OS	Hepatitis A virus (strain 24a).			
OC	Viruses: ssRNA positive-strand viruses, no DNA stage; Picornaviridae;			
OC	Hepatitis A virus.			
OX	NCBI_TaxID=12094;			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE-91162758; PubMed-1705995;			
RA	Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstein S.M.,			
RA	Cromeans T., Jansen R.W.;			
RT	Antigenic and genetic variation in cytopathic hepatitis A virus			
RT	variants arising during persistent infection: evidence for genetic			
RT	recombination.;			
RL	J. Virol. 65:2056-2065(1991).			
CC	- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +			
CC	[RNA](N).			
CC	- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,			
CC	EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,			
CC	VP3, AND VP4.			
CC	- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.			
CC	- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			
DR	EMBL: M59810; AAA5468.1; -			
DR	MEROPS: C03.005; -			
DR	InterPro: IPR004004; Calic1_POL_hel.			
DR	InterPro: IPR000605; RNA_helicase.			
DR	InterPro: IPR007095; RNA_POL_DS_PS.			
DR	InterPro: IPR001205; RNA_POL_P3D.			
DR	InterPro: IPR007094; RNA_POL_PSV1r.			
DR	Pfam: PF00680; RNA_dep_RNA_pol.1.			
DR	Pfam: PF00910; RNA_helicase.1.			
DR	PRINTS: PR00918; CALICVIRUSNS.			
KW	POLYPROTEIN; Coat protein; Core protein; Transferrase;			
KW	RNA-directed RNA polymerase; Hydrolyase; Thiol protease.			
FT	CHAIN 1 23			
FT	CHAIN 24 245			
FT	CHAIN 246 491			
FT	CHAIN 492 794			
FT	CHAIN 795 900			
FT	CHAIN CORE PROTEIN VP4 (P1A).			
FT	CHAIN COAT PROTEIN VP3 (P1B).			
FT	CHAIN COAT PROTEIN VP3 (P1C).			
FT	CHAIN COAT PROTEIN VP1 (P1D).			
FT	CHAIN CORE PROTEIN P2A.			

FT CHAIN 901 1087 CORE PROTEIN P2B.  
FT CHAIN 1088 1422 CORE PROTEIN P2C.  
FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.  
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.  
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.  
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.  
SQ SEQUENCE 2226 AA; 251152 MW; 6CD85A91DB84E2BF CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2226;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
|||||  
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

## RESULT 2

POLG\_HPAV4 STANDARD; PRT; 2226 AA.  
ID POLG\_HPAV4  
AC P26581;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DE 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
DE P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain 43c).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
CX NCBI\_TaxID=12095;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91162758; PubMed=1705995;  
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,  
RA Crommons T., Jansen R.W.;  
RT "Antigenic and genetic variation in cytopathic hepatitis A virus  
RT variants arising during persistent infection: evidence for genetic  
RT recombination.";  
RL J. Virol. 65:2056-2065(1991).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA}(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: M59809; AAA45469.1; -.  
DR MEROPS: C03.005; -.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR007094; RNA\_pol\_PSVir.  
DR Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam: PF00910; RNA\_helicase; 1.  
KW Polypeptide; Coat protein; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 794  
FT CHAIN 795 900  
FT CHAIN 901 1087  
FT CHAIN 1088 1422  
CORE PROTEIN P2A.  
CORE PROTEIN P2B.  
CORE PROTEIN P2C.

FT CHAIN 1423 1495 PROBABLE PROTEIN P3A.  
FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.  
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.  
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.  
SQ SEQUENCE 2226 AA; 251107 MW; 403B4CA80809BE75 CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2226;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
|||||  
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

## RESULT 3

POLG\_HPAV8 STANDARD; PRT; 2226 AA.  
ID POLG\_HPAV8  
AC P26582;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DE 15-SEP-2003 (Rel. 42, Last annotation update)  
DE Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
DE P3D (EC 2.7.7.48)].  
OS Hepatitis A virus (strain 18f).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
CX NCBI\_TaxID=12096;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91162758; PubMed=1705995;  
RA Lemon S.M., Murphy P.C., Shields P.A., Ping L.H., Feinstone S.M.,  
RA Crommons T., Jansen R.W.;  
RT "Antigenic and genetic variation in cytopathic hepatitis A virus  
RT variants arising during persistent infection: evidence for genetic  
RT recombination.";  
RL J. Virol. 65:2056-2065(1991).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA}(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHERAL UNITS,  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: M59608; AAA45467.1; -.  
DR PDB: 1QAT; 1S-MAY-00.  
DR MEROPS: C03.005; -.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR007094; RNA\_pol\_PSVir.  
DR Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam: PF00910; RNA\_helicase; 1.  
KW Polypeptide; Coat protein; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 794  
FT CHAIN 795 900  
FT CHAIN 901 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1495  
CORE PROTEIN P2A.  
CORE PROTEIN P2B.  
CORE PROTEIN P2C.  
PROBABLE PROTEIN P3A.

```

FT CHAIN 1496 1518 PROBABLE PROTEIN P3B.
FT CHAIN 1519 1737 PROBABLE PROTEIN P3C.
FT CHAIN 1738 2226 RNA-DIRECTED POLYMERASE P3D.
SQ SEQUENCE 2226 AA; 251292 MW; 24964463396C8DBB CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2226;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 4
POLG_HPAVL STANDARD; PRT; 2227 AA.
AC P06417; P06443; Q81082;
DT 01-AUG-1988 (rel. 08; Created)
01-AUG-1988 (rel. 08; Last sequence update)
15-SEP-2003 (rel. 42; Last annotation update)
Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain HM-175).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Wild type;
RX MEDLINE=87061253; PubMed=3023706;
RA Cohen J.I., Ticehurst J.R., Purcell R.H., Buckler-White A.,
Baroudy B.M.;
RT *Complete nucleotide sequence of wild-type hepatitis A virus:
RT comparison with different strains of hepatitis A virus and other
RT picornaviruses.*
RL J. Virol. 61:50-59(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Attenuated;
RX MEDLINE=87175701; PubMed=3031686;
RA Cohen J.I., Rosenblum B., Ticehurst J.R., Deemer R.J., Feinstone S.M.,
Purcell R.H.;
RT *Complete nucleotide sequence of an attenuated hepatitis A virus:
RT comparison with wild-type virus.*
RL Proc. Natl. Acad. Sci. U.S.A. 84:2497-2501(1987).
RN [3]
RP SEQUENCE OF 1-854 AND 1724-2227 FROM N.A.
RX MEDLINE=85166289; PubMed=2984684;
RA Baroudy B.M., Ticehurst J.R., Mele T.A., Maizel J.V. Jr.,
Purcell R.H., Feinstone S.M.;
RT *Sequence analysis of hepatitis A virus cDNA coding for capsid
RT proteins and RNA polymerase.*
RL Proc. Natl. Acad. Sci. U.S.A. 82:2143-2147(1985).
RN [4]
RP CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
RN [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PIV: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC FROM HM-175 BY 32 PASSAGES IN AFRICAN GREEN MONKEY KIDNEY CELLS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -1- CAUTION: REF.3 SEQUENCE HAS A NUMBER OF CONFLICTS WITH THAT
CC SHOWN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M14114; AAA5475.1; -
DR EMBL: M14707; AAA5465.1; -
DR EMBL: M14707; AAA5466.1; ALT_INIT.
DR EMBL: M16532; AAA5471.1; -
DR PIR: A25981; GNNYMK.
DR PIR: A94149; GNNYMK.
DR PDB: 1HAV; 23-DEC-96.
DR MEROPS: C03.005; -.
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PR00680; RNA_dep_RNA_pol; 1.
DR Pfam: PR00910; RNA_helicase; 1.
DR PRINTS: PR009148; CALICVIRUSNS.
KW Polypeptide; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolyase; Thiol protease; 3D-structure.
FT CHAIN 1 245
FT CHAIN 23
FT CHAIN 245
FT CHAIN 245
FT CHAIN 491
FT CHAIN 836
FT CHAIN 837
FT CHAIN 980
FT CHAIN 981
FT CHAIN 1087
FT CHAIN 1088
FT CHAIN 1422
FT CHAIN 1423
FT CHAIN 1497
FT CHAIN 1519
FT CHAIN 1520
FT CHAIN 1738
FT CHAIN 1739
FT CHAIN 2227
FT CHAIN 777
FT CHAIN 764
FT CHAIN 821
FT CHAIN 1052
FT CHAIN 1052
FT CHAIN 1062
FT CHAIN 1118
FT CHAIN 1151
FT CHAIN 1151
FT CHAIN 1163
FT CHAIN 1277
FT CHAIN 1277
FT CHAIN 1500
FT CHAIN 1500
FT CHAIN 1805
FT CHAIN 1805
FT CHAIN 1930
SQ SEQUENCE 2227 AA; 251506 MW; 01E225E7AE8740A6 CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2227;
Best Local Similarity 100.0%; Pred. No. 1.9e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVQEIKEQ 25
DB 956 WLNPKKINLADRMGLSGVQEIKEQ 980

RESULT 5
POLG_HPAVL STANDARD; PRT; 2227 AA.
AC P06441;
DT 01-JAN-1988 (rel. 06; Created)
01-JAN-1988 (rel. 06; Last sequence update)
28-FEB-2003 (rel. 41; Last annotation update)
Genome polypeptide [Contains: Coat proteins VP1 TO VP4; Core proteins
P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase
P3D (EC 2.7.7.48)].
DE Hepatitis A virus (strain 1A).
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Hepatovirus.
OX NCBI_TaxID=12099;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85190549; PubMed=2986127;
RA Najarian R., Caput D., Gee W.W., Potter S.J., Renard A.,
Merryweather J., van Nest G., Dina D.;

```

RT "Primary structure and gene organization of human hepatitis A virus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 82:2627-2631(1985).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +  
CC (RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS.  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: K02990; AAA45472.1; -.  
DR PIR: A03903; GNNYHR.  
DR MEROPS: C03.005; -.  
DR InterPro: IPR004004; Calicel.pol.hel.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA.pol.DS\_PS.  
DR InterPro: IPR001205; RNA.pol.P3D.  
DR InterPro: IPR007094; RNA.pol.PSVlr.  
DR Pfam: PF00680; RNA\_dep.RNA\_pol; 1.  
DR Pfam: PF00910; RNA\_helicase; 1.  
DR PRINIS: PR00918; CALICIVIRUSNS.  
DR Polyprotein: Coat protein; Core protein; Transferase;  
KM RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 836  
FT CHAIN 837 980  
FT CHAIN 981 1076  
FT CHAIN 1077 1422  
FT CHAIN 1423 1484  
FT CHAIN 1485 1507  
FT CHAIN 1508 1678  
FT CHAIN 1679 2227  
SO SEQUENCE 2227 AA; 251898 MW; 99A7354B4CD2799C CRC64;  
Query Match 100.0%; Score 129; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 WLNPKNINLADRMGLSGVOEIKQ 25  
|||||  
956 WLNPKNINLADRMGLSGVOEIKQ 980

RT "The entire nucleotide sequence of the genome of human hepatitis A  
RT virus (isolate MBH).";  
RL Virus Res. 8:153-171(1987).  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate +  
CC (RNA)(N).  
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS.  
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
CC VP3, AND VP4.  
CC -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
CC -1- SIMILARITY: THIS VIRUS WAS ISOLATED FROM THE STOOL OF A PATIENT  
CC WHO HAD BEEN INFECTED IN THE AFRICAN MEDITERRANEAN REGION.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M20273; AAA45474.1; -.  
DR MEROPS: C03.005; -.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA.pol.DS\_PS.  
DR InterPro: IPR001205; RNA.pol.P3D.  
DR InterPro: IPR007094; RNA.pol.PSVlr.  
DR Pfam: PF00680; RNA\_dep.RNA\_pol; 1.  
DR Pfam: PF00910; RNA\_helicase; 1.  
DR Polyprotein: Coat protein; Core protein; Transferase;  
KM RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
FT CHAIN 1 23  
FT CHAIN 24 245  
FT CHAIN 246 491  
FT CHAIN 492 836  
FT CHAIN 837 980  
FT CHAIN 981 1087  
FT CHAIN 1088 1422  
FT CHAIN 1423 1496  
FT CHAIN 1497 1519  
FT CHAIN 1520 1738  
FT CHAIN 1739 2227  
SO SEQUENCE 2227 AA; 251425 MW; EC983ED2A7C86349 CRC64;  
Query Match 100.0%; Score 129; DB 1; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
1 WLNPKNINLADRMGLSGVOEIKQ 25  
|||||  
956 WLNPKNINLADRMGLSGVOEIKQ 980

RESULT 7  
POLG\_HPAVS STANDARD; PRT: 2230 AA.  
ID POLG\_HPAVS  
AC P14553;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polyprotein [contains: Coat proteins VP1 TO VP4; Core proteins  
DE P2A TO P2C; Probable proteins P3A TO P3C; RNA-directed RNA polymerase  
DE P3D (EC 2.7.7.48)].  
OS Simian hepatitis A virus (strain AGM-27).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID-12102;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-91311420; PubMed-1649901;  
RA Tsarev S.A., Emerson S.U., Balaayan M.S., Ticehurst J.R.,  
RA Purcell R.H.;  
RT "Simian hepatitis A virus (HAV) strain AGM-27: comparison of genome

```

RT structure and growth in cell culture with other HAV strains."
RL J. Gen. Virol. 72:1677-1683(1991).
(12)
RP SEQUENCE OF 1750-2164 FROM N.A.
RX MEDLINE-89232168; PubMed-2541023;
RA Balayan M.S., Kusov Y.Y., Andjaparidze A.G., Tsarev S.A.,
RA Vardolov E.D., Chizhikov V.E., Blinov V.M., Vasilenko S.K.;
RT Variations in genome fragments coding for RNA polymerase in human
RT and simian hepatitis A viruses.*
RL FEBS Lett. 247:425-428(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC [RNA](N).
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00924; BAA00766.1;
DR EMBL: X15461; CAA33490.1;
DR PIR: A30470; GNNYSA.
DR MEROPS: C03.005;
DR InterPro: IPR004004; Calic1_pol_hel.
DR InterPro: IPR000605; RNA_helicase.
DR InterPro: IPR001095; RNA_pol_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00680; RNA_dep_RNA_pol.1.
DR Pfam: PF00910; RNA_helicase.1.
DR PRINTS: PR00918; CALICVIRUSNS.
DR PolyProtein: Coat protein; Core protein; Transferase;
DR RNA-directed RNA polymerase; Hydrolase; Thiol protease.
KW CHAIN 1 27
FT CHAIN 28 249 COAT PROTEIN VP4 (PIA).
FT CHAIN 250 495 COAT PROTEIN VP2 (PIB).
FT CHAIN 496 795 COAT PROTEIN VP3 (PIC).
FT CHAIN 796 984 COAT PROTEIN VP1 (PID).
FT CHAIN 985 1091 CORE PROTEIN P2A.
FT CHAIN 1092 1426 CORE PROTEIN P2B.
FT CHAIN 1427 1498 CORE PROTEIN P2C.
FT CHAIN 1499 1521 PROBABLE PROTEIN 3A.
FT CHAIN 1522 1741 PROBABLE PROTEIN 3B.
FT CHAIN 1742 2230 PROBABLE PROTEIN 3C.
FT CHAIN 2230 251296 RNA-DIRECTED POLYMERASE 3D.
FT CHAIN 251296 87B3230E324E1F19 CRC64;
SQ SEQUENCE 2230 AA; 251296 MW; 87B3230E324E1F19 CRC64;

Query Match 100.0%; Score 129; DB 1; Length 2230;
Best Local Similarity 100.0%; Pred. No. 1; 19e-11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLPKKNLADRMGLSGVOEIKEQ 25
DB 960 WLPKKNLADRMGLSGVOEIKEQ 984

RESULT 8
HS20_NIPBR STANDARD: PRT; 172 AA.
AC 007160:
DT 01-OCT-1994 (rel. 30, Created)
DT 01-OCT-1994 (rel. 30, Last sequence update)
DT 01-OCT-1994 (rel. 30, Last annotation update)
DE Heat shock protein homolog (HSP20).
GN HSP20.
OS Nipostromylius brasiliensis.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Strongylida;

```

```

OC Trichostromyloidea; Heligmoneillidae; Nipostromyliinae;
OC Nipostromylius.
OX NCBI_TaxID-27835;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE-94081860; PubMed-8259127;
RA Tweedie S., Grigg M.E., Ingram L., Selkirk M.E.;
RT The expression of a small heat shock protein homologue is
RT developmentally regulated in Nipostromylius brasiliensis.*
RL Mol. Biochem. Parasitol. 61:149-154(1993).
CC -1- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X71663; CAA50655.1;
DR PIR: S33416; S33416.
DR InterPro: IPR001436; Crystallin_alpha.
DR InterPro: IPR002068; Hsp20.
DR Pfam: PF00011; HSP20.1.
DR PRINTS: PR00299; ACRYSTALLIN.
DR PROSITE: PS01031; HSP20.1.
KW Heat shock.
SQ SEQUENCE 172 AA; 20227 MW; 2CDAA711CE60B1C0 CRC64;

Query Match 38.8%; Score 50; DB 1; Length 172;
Best Local Similarity 43.5%; Pred. No. 1.9;
Matches 10; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 3 NPKKNLADRMGLSGVOEIKEQ 25
DB 77 NPKKNLADRMGLSGVOEIKEQ 99

RESULT 9
RIR2_ASFW2 STANDARD: PRT; 327 AA.
AC P26713;
DT 01-AUG-1992 (rel. 23, Created)
DT 01-AUG-1992 (rel. 23, Last sequence update)
DT 01-AUG-1992 (rel. 23, Last annotation update)
DE Ribonucleoside-diphosphate reductase small chain (RC 1.17.4.1)
DE (Ribonucleotide reductase).
OS African swine fever virus (isolate Malawi L1 20/1) (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfarvirus.
OX NCBI_TaxID-10500;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE-91335775; PubMed-1871976;
RA Bousnell M., Shaw K., Yanez R.J., Vinuela E., Dixon L.;
RT The sequences of the ribonucleotide reductase genes from African
RT swine fever virus show considerable homology with those of the
RT orthopoxvirus, vaccinia virus.*
RL Virology 184:411-416(1991).
CC -1- FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -1- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thiorodoxin + H(2)O -> ribonucleoside diphosphate + reduced
CC thiorodoxin.
CC -1- COFACTOR: BINDS 2 IRON IONS.
CC -1- PATHWAY: DNA replication pathway; first step.
CC -1- SUBUNIT: Heterodimer of a large and a small chain.
CC -1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC SMALL CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL, M64728; -, NOT_ANNOTATED_CDS.
CC
CC DR PIR, B40568; RDVZAS.
CC DR HSSP, P11157; 1XSM.
CC DR InterPro, IPR000358; RibonucL_redctse.
CC DR Pfam, PF00268; ribonuc_red.sm.1.
CC DR PROSITE, PS00368; RIBORED_SMALL; 1.
CC
CC KM Oxidoreductase; DNA replication; Iron.
CC FT METAL 70
CC FT METAL 101 IRON 1 (BY SIMILARITY).
CC FT METAL 104 IRON 1 AND 2 (BY SIMILARITY).
CC FT METAL 164 IRON 1 (BY SIMILARITY).
CC FT METAL 198 IRON 2 (BY SIMILARITY).
CC FT METAL 201 IRON 2 (BY SIMILARITY).
CC FT ACT_SITE 108 IRON 2 (BY SIMILARITY).
CC FT SEQUENCE 327 AA; 38966 MW; E78508DB1978F4B0 CRC64;
CC
CC Query Match 38.0%; Score 49; DB 1; Length 327;
CC Best Local Similarity 31.8%; Pred. No. 5.2;
CC Matches 7; Conservative 9; Mismatches 6; Indels 0; Gaps 0;
CC
CC 1 WLNPKKINMADRMGLSGVQEI 22
CC ::::::::::::::::::::
CC Db 145 WMDPARNSLGERLVGFAAVEGI 166
CC
CC RESULT 10
CC POP6_YEAST STANDARD; PRT; 158 AA.
CC
CC AC P53218;
CC DT 01-OCT-1996 (Rel. 34, Created)
CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Ribonucleases P/MRP protein subunit POP6 (Ec 3.1.26.5) (RNases P/MRP
CC 18.2 kDa subunit) (RNA processing protein POP6).
CC GN POP6 OR YGR030C.
CC OS Saccharomyces cerevisiae (Baker's yeast).
CC CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CC CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
CC CX NCBI_TaxID=4932;
CC
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=5288C;
CC RX MEDLINE=97435481; PubMed=9290212;
CC "Rieger M., Brueckner M., Schaefer M., Mueller-Ruer S.;
CC "Sequence analysis of 203 kilobases from Saccharomyces cerevisiae
CC chromosome VII."
CC RL Yeast 13:1077-1090(1997).
CC
CC -1 FUNCTION: COMPONENT OF RIBONUCLEASE P, A PROTEIN COMPLEX THAT
CC GENERATES MATURE tRNA MOLECULES BY CLEAVING THEIR 5' ENDS.
CC ALSO A COMPONENT OF RNASE MRP.
CC -1 CATALYTIC ACTIVITY: Endonucleolytic cleavage of RNA, removing 5'-
CC extra-nucleotide from tRNA precursor.
CC -1 SUBUNIT: COMPONENT OF NUCLEAR RNASE P AND RNASE MRP RNASE P
CC RIBONUCLEOPROTEINS. RNASE P CONSISTS OF A RNA MOLETY AND AT LEAST
CC 8 PROTEIN SUBUNTS; POP1, POP3, POP4, POP5, POP6, POP7, POP8 AND
CC REP1.
CC -1 SUBCELLULAR LOCATION: Nuclear (Potential).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL, Z72815; CAAG7018.1; -.
CC PIR, S64321; S64321.

```

```

DR SGD: S0003262; POP6.
DR GO: GO:0000172; C:ribonuclease mitochondrial RNA processing c. . ; IDA.
DR GO: GO:0005655; C:ribonuclease p complex; IDA.
DR GO: GO:0000171; F:ribonuclease MP activity; IDA.
DR GO: GO:0004526; F:ribonuclease P activity; IDA.
FW Hydrolase; Nuclear protein; tRNA processing; Coiled coil.
KW DOMAIN
CC COILED COIL (POTENTIAL)
SO SEQUENCE 158 AA; 18210 MW; 6C27A73FADF51181 CRC64;

Query Match 37.2%; Score 48; DB 1; Length 158;
Best Local Similarity 66.7%; Pred. No. 3.4;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 5 KKLADRMGLSGV 19
Db 64 KOINMADRSLGQCV 78
|:::|||||
RESULT 11
RIR2_ASF87 STANDARD; PRT; 334 AA.
ID RIR2_ASF87 STANDARD; PRT; 334 AA.
AC P42492;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
DE (Ribonucleotide reductase).
GN F334L.
OS African swine fever virus (strain BA71V) (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxId=10498;
RN [1]
RP SEQUENCE FROM N.A.
RA yanez R.J., Rodriguez J.M., Nogal M.L., Yuste L., Enriquez C.,
RT Rodriguez J.F., Vinuela E.;
RT Analysis of the complete nucleotide sequence of African swine fever
virus. ";
RL Virology 208:249-278(1995).
CC -1 FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS. *
CC -1 CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thioedoxin + H(2)O = ribonucleoside diphosphate + reduced
CC thioedoxin.
CC -1 COFACTOR: BINDS 2 IRON IONS (BY SIMILARITY).
CC -1 PATHWAY: DNA replication pathway; first step.
CC -1 SUBUNIT: Heterodimer of a large and a small chain.
CC -1 SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE
CC SMALL CHAIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: U18466; AAA65274.1; -.
DR HSSP: P11157; IXSM.
DR InterPro: IPRO00358; RibonucL_reductse.
DR Pfam: PF00268; ribonuc_red_sm.1.
DR PROSITE: PS00368; RIBORED_SMALL; 1.
KW Oxidoreductase; DNA replication; Iron.
FT METAL 77 IRON 1 (BY SIMILARITY).
FT METAL 108 IRON 1 AND 2 (BY SIMILARITY).
FT METAL 111 IRON 1 (BY SIMILARITY).
FT METAL 171 IRON 2 (BY SIMILARITY).
FT METAL 205 IRON 2 (BY SIMILARITY).
FT METAL 208 IRON 2 (BY SIMILARITY).
FT ACT_SITE 115 BY SIMILARITY.
SQ SEQUENCE 334 AA; 39806 MW; 9501955EFD0719CF CRC64;

Query Match 37.2%; Score 48; DB 1; Length 334;
Best Local Similarity 31.8%; Pred. No. 7.5;
```



Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIK 22  
DB 152 WMDPRNTLGERLVGFAAVEGI 173

RESULT 12  
LEPA\_STRCO STANDARD: PRT: 622 AA.  
AC 09RDC9;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-SEP-2003 (Rel. 42, Last annotation update)  
DE GTP-binding protein lepa.  
GN LEPA OR SC02562 OR SCC77.29C.  
OS Streptomyces coelicolor.  
OC Bacteria; Actinobacteria; Actinomycetales;  
OC Streptomycinae; Streptomycetaceae; Streptomyces.  
NCBI\_TaxID=1902;

SEQUENCE FROM N.A.  
RC STRAIN-A3(2) / M145;  
RA MEDLINE-21996410; PubMed-12000953;  
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
RA Huang C.-H., Kleser T., Larke L., Murphy J., Oliver K., O'Neill S.,  
RA Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,  
RA Seeger K., Saunders D., Sharp S., Squares R., Taylor K.,  
RA Warren T., Watzorek A., Woodward J., Barrell B.G., Parkhill J.,  
RA Hopwood D.A.;  
RT \*Complete genome sequence of the model actinomycete Streptomyces  
RT coelicolor A3(2).  
RL Nature 417:141-147(2002).  
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
LEPA SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: AL939113; CAB66240.1; -  
HSP: P13551; 2EPG.  
HAMAP: MF\_00071; 1.  
DR InterPro: IPR000795; EF\_GTPbind.  
DR InterPro: IPR000640; EFG\_C.  
DR InterPro: IPR004161; EFTU\_D2.  
DR InterPro: IPR006297; Lepa.  
DR InterPro: IPR005225; Small\_GTP.  
DR Pfam: PF00679; EFG\_C.1.  
DR Pfam: PF00009; GTP\_EFTU.1.  
DR Pfam: PF03144; GTP\_EFTU\_D2.1.  
DR PRINTS: PRO0315; ELONGATNFC.  
DR TIGRFAMs: TIGR01393; lepa.1.  
DR TIGRFAMs: TIGR00231; small\_GTP.1.  
DR PROSITE: PS00301; EFACITOR\_GTP.1.  
KW GTP-binding; Complete proteome.  
FT NP\_BIND 26 33 GTP (BY SIMILARITY).  
FT NP\_BIND 94 98 GTP (BY SIMILARITY).  
FT NP\_BIND 148 151 GTP (BY SIMILARITY).  
SQ SEQUENCE 622 AA: 68378 MW: 835076FA2A80C7C CRC64:

Query Match 37.2%; Score 48; DB 1; Length 622;  
Best Local Similarity 57.9%; Pred. No. 14;  
Matches 11; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 6 KINLADRMGLSGVOEIK 24

DB 32 KSTLADRMQLTGVGEORQ 50

RESULT 13  
TPR\_PORGI STANDARD: PRT: 482 AA.  
AC P25806;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Thiol protease (EC 3.4.22.-).  
GN TPR.  
OS Porphyromonas gingivalis (Bacteroides gingivalis).  
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;  
OC Porphyromonadaceae; Porphyromonas.  
NCBI\_TaxID=837;

SEQUENCE FROM N.A.  
RP STRAIN-W83;  
RC MEDLINE-92347990; PubMed-1322368;  
RA Bourgeois G., Lapointe H., Pelouquin P., Mayrand D.;  
RT "Cloning, expression, and sequencing of a protease gene (tpr) from  
RT Porphyromonas gingivalis W83 in Escherichia coli.";  
RL Infect. Immun. 60:3186-3192(1992).  
CC -1- FUNCTION: THIOL PROTEASE. PROBABLY AN IMPORTANT VIRULENCE FACTOR.  
CC -1- ENZYME REGULATION: INACTIVE BELOW 20 DEGREES CELSIUS AND PH 6.0.  
CC INHIBITED BY DIVALENT CATIONS.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C1.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: M84471; AAA25652.1; -  
DR PIR: S27608; S27608.  
DR MEROPS: C01.090; -  
DR InterPro: IPR001300; Protease\_C2.  
DR InterPro: IPR000169; SHProl\_acsite.  
DR Pfam: PF00648; Peptidase\_C2.1.  
DR SMART: SM00230; Cyspc.1.  
DR PROSITE: PS00639; THIOL\_PROTEASE\_HIS; FALSE\_NEG.  
DR PROSITE: PS00640; THIOL\_PROTEASE\_ASN; FALSE\_NEG.  
DR PROSITE: PS00139; THIOL\_PROTEASE\_CYS.1.  
KW Hydrolyase; Thiol protease; Virulence.  
FT ACT\_SITE 229 229 BY SIMILARITY.  
FT ACT\_SITE 407 407 BY SIMILARITY.  
FT ACT\_SITE 427 427 BY SIMILARITY.  
SQ SEQUENCE 482 AA: 55104 MW: 7E1E0788C2A70A50 CRC64:

Query Match 36.8%; Score 47.5; DB 1; Length 482;  
Best Local Similarity 33.3%; Pred. No. 13;  
Matches 8; Conservative 9; Mismatches 6; Indels 1; Gaps 1;

OY 1 WLNPKKINLADRMGLSGVOEIK 24  
DB 80 WSNPRTTIVRYLG-SSMODLSK 102

RESULT 14  
HS12\_CAEEL STANDARD: PRT: 145 AA.  
AC P06562;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Heat shock protein HSP16-2.  
GN HSP16-2 OR Y46H3A.3.  
OS Caenorhabditis elegans.

```

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;
OX Rhabditiidae; Peloderinae; Caenorhabditis.
RN NCBI_TaxID=6239;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE-86304344; PubMed-3017958;
RA Jones D., Russnak R.H., Kay R.J., Candido E.P.M.;
RT "Structure, expression, and evolution of a heat shock gene locus in
RT Caenorhabditis elegans that is flanked by repetitive elements.";
RL J. Biol. Chem. 261:12006-12015(1986).
RN [2]
RN SEQUENCE FROM N.A.
RP STRAIN-Bristol N2;
RC Stoneking T., Wohlmann P., Lennox S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RN REVISIONS.
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
CC -1 SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20) FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M14334; AAA28071.1; -.
DR EMBL; AC006774; AAF60615.1; -.
DR PIR; B25199; B25199.
DR WormRep; Y46H3A.3; CE22002.
DR InterPro; IPR002068; HSP20.
DR Pfam; PF00011; HSP20_1.
DR PROSITE; PS01031; HSP20_1.
KW Heat shock; Multigene family.
SQ SEQUENCE 145 AA; 16242 MW; 8A73449F99161889 CRC64;
Query Match 36.4%; Score 47; DB 1; Length 145;
Best Local Similarity 55.6%; Pred. No. 4.5;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 6 KINLADRMGLSGVQEIK 23
DQ ||||| | : | : | : | : |
DQ 65 KINLDGRTLSTGGQEELK 82
DQ
DQ ILT 15
DQ 2_CAEEL
AC AC P42170; STANDARD; PRT; 381 AA.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
DE (Ribonucleotide reductase).
DE RNR-2 OR C03C10.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiidae; Rhabditoidea;
OC Rhabditiidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
OX [1]
OX SEQUENCE FROM N.A.
RP STRAIN-Bristol N2;
RA Berks M.;
RL Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
CC -1 FUNCTION: PROVIDES THE PRECURSORS NECESSARY FOR DNA SYNTHESIS.
CC -1 CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC thiorodoxin + H(2)O = ribonucleoside diphosphate + reduced
CC thiorodoxin.
CC -1 COFACTOR: BINDS 2 IRON IONS (BY SIMILARITY).

```

CC	-1- PATHWAY: DNA replication pathway; first step.
CC	-1- SUBUNIT: Heterodimer of a large and a small chain.
CC	-1- SIMILARITY: BELONGS TO THE RIBONUCLEOSIDE DIPHOSPHATE REDUCTASE SMALL CHAIN FAMILY.
CC	-----
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isdb.ch/announce/">http://www.isdb.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL: Z35637; CAN84688.1; .
DR	PIR: T18876; T18876.
DR	HSSP: P1157; IXSM.
DR	WormPeP: C03C10.3; CE00874.
DR	InterPro: IPR000358; RlbonucI_redctse.
KM	Pfam: PF00268; rlbouc_red.sm; 1.
DR	PROSITE: PS00368; RIBORED_SMALL; 1.
KM	Oxidoreductase; DNA replication; Iron.
FT	METAL 130 130 IRON 1 (BY SIMILARITY).
FT	METAL 161 161 IRON 1 AND 2 (BY SIMILARITY).
FT	METAL 164 164 IRON 1 (BY SIMILARITY).
FT	METAL 224 224 IRON 2 (BY SIMILARITY).
FT	METAL 258 258 IRON 2 (BY SIMILARITY).
FT	METAL 261 261 IRON 2 (BY SIMILARITY).
FT	ACT SITE 168 168 BY SIMILARITY.
SQ	SEQUENCE 381 AA; 44289 MW; 75497147ABF36C59 CRC64;
Oy	Query Match 36.4%; Score 47; DB 1; Length 381; Best Local Similarity 31.8%; Fred. No. 12;
Db	Matches 7; Conservative 8; Mismatches 7; Indels 0; Gaps 0;  1 WLNPKRIADRMGLSGVOEI 22   :    : :: : :  : 205 WISDKASFARLLIAFAVEGI 226

Search completed: October 1, 2003, 09:57:57  
Job time : 10.2353 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:48:51 ; Search time 41.4706 Seconds

(without alignments)  
155.564 Million cell updates/sec

Title: US-09-171-432a-48

Perfect score: 129

Sequence: 1 WLNPKINLADRLGLSGVQEIKEQ 25

Scoring table: BLOSUM62

Searched: 830525 seqs, 258052604 residues

al number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SPTREMBL\_23.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mmc.\*

8: sp\_organelle.\*

9: sp\_phage.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriophage.\*

17: sp\_archaea.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	129	100.0	251	12	O9ENP8	O9enp8 hepatitis a
2	129	100.0	251	12	O9ENP9	O9enp9 hepatitis a
3	129	100.0	251	12	O9ENQ4	O9enq4 hepatitis a
4	129	100.0	251	12	O9ENQ2	O9enq2 hepatitis a
5	129	100.0	251	12	O9ENN2	O9enn2 hepatitis a
6	129	100.0	251	12	O9ENP2	O9enp2 hepatitis a
7	129	100.0	251	12	O9ENP6	O9enp6 hepatitis a
8	129	100.0	251	12	O9ENP1	O9enp1 hepatitis a
9	129	100.0	251	12	O9ENP4	O9enp4 hepatitis a
10	129	100.0	251	12	O9ENQ9	O9enq9 hepatitis a
11	129	100.0	251	12	O9ENN4	O9enn4 hepatitis a
12	129	100.0	251	12	O9ENP5	O9enp5 hepatitis a
13	129	100.0	251	12	O9ENQ3	O9enq3 hepatitis a
14	129	100.0	251	12	O9ENP7	O9enp7 hepatitis a
15	129	100.0	251	12	O9ENQ5	O9enq5 hepatitis a
16	129	100.0	251	12	O9ENQ5	O9enq5 hepatitis a

17	129	100.0	251	12	O9ENP9	O9enp9 hepatitis a
18	129	100.0	251	12	O9ENP6	O9enp6 hepatitis a
19	129	100.0	251	12	O9ENP6	O9enp6 hepatitis a
20	129	100.0	251	12	O9ENQ8	O9enq8 hepatitis a
21	129	100.0	251	12	O9ENP3	O9enp3 hepatitis a
22	129	100.0	251	12	O9ENQ0	O9enq0 hepatitis a
23	129	100.0	1124	12	O84780	O84780 hepatitis a
24	129	100.0	1161	12	O05794	O05794 hepatitis a
25	129	100.0	2216	12	O9WMA2	O9wma2 hepatitis a
26	129	100.0	2218	12	O67824	O67824 hepatitis a
27	129	100.0	2218	12	O67817	O67817 hepatitis a
28	129	100.0	2225	12	O9DLJ2	O9dlj2 hepatitis a
29	129	100.0	2227	12	O9WMA0	O9wma0 hepatitis a
30	129	100.0	2227	12	O9WMA3	O9wma3 hepatitis a
31	129	100.0	2227	12	O67825	O67825 hepatitis a
32	129	100.0	2227	12	O9WMA9	O9wma9 hepatitis a
33	129	100.0	2227	12	O8QV03	O8qv03 hepatitis a
34	129	100.0	2227	12	O9WMA1	O9wma1 hepatitis a
35	129	100.0	2227	12	O67826	O67826 hepatitis a
36	129	100.0	2227	12	O8V0N6	O8v0n6 hepatitis a
37	129	100.0	2227	12	O9IFH5	O9ifh5 hepatitis a
38	129	100.0	2227	12	O9WMA4	O9wma4 hepatitis a
39	125	96.9	251	12	O9ENP0	O9enp0 hepatitis a
40	125	96.9	251	12	O9ENN5	O9enn5 hepatitis a
41	125	96.9	251	12	O9ENN7	O9enn7 hepatitis a
42	124	96.1	2225	12	O9DWR1	O9dwr1 hepatitis a
43	124	96.1	2227	12	O8OR16	O8or16 hepatitis a
44	121	93.8	251	12	O9ENR0	O9enr0 hepatitis a
45	121	93.8	251	12	O9ENN8	O9enn8 hepatitis a

## ALIGNMENTS

RESULT 1	
O9ENP8	PRELIMINARY; PRT; 251 AA.
AC O9ENP8;	
DT 01-MAR-2001 (TREMBLrel. 16, Created)	
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	
DE Polyprotein (Fragment).	
OS Hepatitis A virus.	
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;	
OC Hepatovirus.	
OX NCBI_TaxID=12092;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=A304;	
RA Fujiwara K.;	
RT "Hepatitis A virus."	
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.	
DR EMBL: AB047665; BMB12173.1; -.	
FT NON_TER	1
FT NON_TER	251
SQ SEQUENCE	251 AA; 28693 MW; C32AD5651506751A CRC64;
Query Match	100.0%; Score 129; DB 12; Length 251;
Best Local Similarity	100.0%; Pred. No. 9.3e-12;
Matches 25; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WLNPKINLADRLGLSGVQEIKEQ 25	
DB 120 WLNPKINLADRLGLSGVQEIKEQ 144	
RESULT 2	
O9ENN9	PRELIMINARY; PRT; 251 AA.
AC O9ENN9;	
DT 01-MAR-2001 (TREMBLrel. 16, Created)	
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)	

DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A68;  
RA Fujiwara K.;  
RT "hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AB047674; BAB12182.1; -.  
FT NON\_TER 1  
FT SEQUENCE 251 AA; 28666 MW; 5A315F46C830D214 CRC64;  
SQ

Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 120 WLNPKKINLADRMGLSGVOEIKQ 144  
|||||

RESULT 3  
Q9ENQ4 PRELIMINARY; PRT; 251 AA.  
ID Q9ENQ4;  
AC Q9ENQ4;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A201;  
RA Fujiwara K.;  
RT "hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AB047659; BAB12167.1; -.  
FT NON\_TER 1  
FT SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;  
SQ

Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 120 WLNPKKINLADRMGLSGVOEIKQ 144  
|||||

RESULT 4  
Q9ENQ2 PRELIMINARY; PRT; 251 AA.  
ID Q9ENQ2;  
AC Q9ENQ2;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A205;  
RA Fujiwara K.;

RT "hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AB047661; BAB12169.1; -.  
FT NON\_TER 1  
FT SEQUENCE 251 AA; 28723 MW; 558A2D3664C7343C CRC64;  
SQ

Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 120 WLNPKKINLADRMGLSGVOEIKQ 144  
|||||

Query 5  
Q9ENN2 PRELIMINARY; PRT; 251 AA.  
ID Q9ENN2;  
AC Q9ENN2;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A9;  
RA Fujiwara K.;  
RT "hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AB047681; BAB12189.1; -.  
FT NON\_TER 1  
FT SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;  
SQ

Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 120 WLNPKKINLADRMGLSGVOEIKQ 144  
|||||

Query 6  
Q9ENP2 PRELIMINARY; PRT; 251 AA.  
ID Q9ENP2;  
AC Q9ENP2;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A5;  
RA Fujiwara K.;  
RT "hepatitis A virus.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AB047671; BAB12179.1; -.  
FT NON\_TER 1  
FT SEQUENCE 251 AA; 28622 MW; B55C3CD146D39D02 CRC64;  
SQ

Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
|||||  
DB 120 WLNPKKINLADRMGLSGVOEIKQ 144

## RESULT 7

O9EN06 PRELIMINARY; PRT; 251 AA.  
AC O9EN06: 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBL\_TaxID-12092;  
RN [1]  
SEQUENCE FROM N.A.  
RA Fujikawa K.;  
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047657; BAB12165.1; -.  
FT NON\_TER 1 251  
SQ SEQUENCE 251 AA; 28699 MW; 56AA86DF60F6D3B CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
|||||  
DB 120 WLNPKKINLADRMGLSGVOEIKQ 144

## RESULT 8

O9ENR1 PRELIMINARY; PRT; 251 AA.  
AC O9ENR1: 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBL\_TaxID-12092;  
RN [1]  
SEQUENCE FROM N.A.  
RA Fujikawa K.;  
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047652; BAB12160.1; -.  
FT NON\_TER 1 251  
SQ SEQUENCE 251 AA; 28749 MW; 58A520D873893445 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
|||||  
DB 120 WLNPKKINLADRMGLSGVOEIKQ 144

## RESULT 9

O9ENP1

ID O9ENP1 PRELIMINARY; PRT; 251 AA.

AC O9ENP1: 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBL\_TaxID-12092;  
RN [1]  
SEQUENCE FROM N.A.  
RA Fujikawa K.;  
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047672; BAB12180.1; -.  
FT NON\_TER 1 251  
SQ SEQUENCE 251 AA; 28663 MW; C7EA6FBD19A1619 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
|||||  
DB 120 WLNPKKINLADRMGLSGVOEIKQ 144

## RESULT 10

O9ENP4 PRELIMINARY; PRT; 251 AA.  
AC O9ENP4: 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBL\_TaxID-12092;  
RN [1]  
SEQUENCE FROM N.A.  
RA Fujikawa K.;  
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047669; BAB12177.1; -.  
FT NON\_TER 1 251  
SQ SEQUENCE 251 AA; 28666 MW; 5A315F46C830D214 CRC64;

Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
|||||  
DB 120 WLNPKKINLADRMGLSGVOEIKQ 144

## RESULT 11

O9EN09 PRELIMINARY; PRT; 251 AA.  
AC O9EN09: 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)  
DE Polypeptide (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.

OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A159;  
RA Fujiwara K.;  
RT "hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047654; BAB12162.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;  
  
Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
Db 120 WLNPKKINLADRMGLSGVOEIKQ 144  
|||||  
SOUT 12  
Q9ENM4 PRELIMINARY; PRT; 251 AA.  
ID Q9ENM4  
AC Q9ENM4  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A77;  
RA Fujiwara K.;  
RT "hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047679; BAB12187.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 251 AA; 28734 MW; 3B96B782882F19D9 CRC64;  
  
Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
Db 120 WLNPKKINLADRMGLSGVOEIKQ 144  
|||||  
RESULT 13  
Q9ENP5 PRELIMINARY; PRT; 251 AA.  
ID Q9ENP5  
AC Q9ENP5  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A407;  
RA Fujiwara K.;  
RT "hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047668; BAB12176.1; -.  
FT NON\_TER 1

FT NON\_TER 251  
SQ SEQUENCE 251 AA; 28614 MW; 8334EF179C757A6D CRC64;  
  
Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
Db 120 WLNPKKINLADRMGLSGVOEIKQ 144  
|||||  
RESULT 14  
Q9ENQ3 PRELIMINARY; PRT; 251 AA.  
ID Q9ENQ3  
AC Q9ENQ3  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A204;  
RA Fujiwara K.;  
RT "hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047660; BAB12168.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 251 AA; 28693 MW; C32AD5651506751A CRC64;  
  
Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
Db 120 WLNPKKINLADRMGLSGVOEIKQ 144  
|||||  
RESULT 15  
Q9ENP7 PRELIMINARY; PRT; 251 AA.  
ID Q9ENP7  
AC Q9ENP7  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis A virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Hepatovirus.  
OX NCBI\_TaxID=12092;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A306;  
RA Fujiwara K.;  
RT "hepatitis A virus."  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB047666; BAB12174.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 251 AA; 28720 MW; C3342482882F19CA CRC64;  
  
Query Match 100.0%; Score 129; DB 12; Length 251;  
Best Local Similarity 100.0%; Pred. No. 9.3e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
|||||

Db 120 WLNPKKINLADRMGLSGVQEIKEQ 144

Search completed: October 1, 2003, 10:02:41  
Job time : 41.4706 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 09:50:25 ; Search time 15.8824 Seconds  
(without alignments)  
66.600 Million cell updates/sec

Title: US-09-171-432A-48  
Perfect score: 129  
Sequence: 1 WLNPKKINLADRMGLSGVOEIKEQ 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

al number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCRTS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	2227	3	US-08-475-886-2
2	129	100.0	2227	3	US-08-475-886-4
3	129	100.0	2227	3	US-08-475-886-6
4	129	100.0	2227	3	US-08-397-232-2
5	129	100.0	2227	3	US-08-397-232-4
6	129	100.0	2227	3	US-09-171-387-2
7	129	100.0	2227	4	US-09-653-499-2
8	129	100.0	2227	4	US-09-653-499-4
9	129	100.0	2227	4	US-09-653-499-6
10	49	38.0	229	3	US-09-248-528-17
11	49	38.0	229	3	US-09-549-108-17
12	49	38.0	229	3	US-09-549-111-17
13	49	38.0	229	3	US-09-549-106-17
14	49	38.0	229	3	US-09-550-394-17
15	47.5	36.8	699	4	US-09-252-991A-17077
16	46	34.9	877	4	US-09-165-396-5
17	45	34.9	877	4	US-09-134-001C-4798
18	44	34.1	495	4	US-09-107-532A-6687
19	42.5	32.9	393	4	US-09-198-452A-565
20	42	32.6	183	4	US-09-107-532A-5935
21	42	32.6	185	2	US-08-691-814B-12
22	42	32.6	185	4	US-09-250-609-11
23	42	32.6	185	4	US-09-250-611-11
24	42	32.6	271	4	US-09-252-991A-21522
25	42	32.6	343	4	US-09-252-991A-26240
26	42	32.6	514	3	US-09-385-028-5
27	42	32.6	514	4	US-09-726-614-5

28	42	32.6	638	2	US-08-846-762-95	Sequence 95, Appl
29	42	32.6	1250	1	US-08-441-139-9	Sequence 9, Appl
30	42	32.6	1404	4	US-08-801-308-1	Sequence 1, Appl
31	41	31.8	166	4	US-09-732-210-1606	Sequence 1606, Ap
32	41	31.8	184	2	US-08-715-204-5	Sequence 5, Appl
33	41	31.8	184	2	US-08-691-814B-50	Sequence 50, Appl
34	41	31.8	184	3	US-09-162-597-5	Sequence 13, Appl
35	41	31.8	184	4	US-09-250-609-13	Sequence 13, Appl
36	41	31.8	184	4	US-09-250-611-13	Sequence 6, Appl
37	41	31.8	366	4	US-08-746-883-6	Sequence 17, Appl
38	41	31.8	370	2	US-08-837-593-7	Sequence 7, Appl
39	41	31.8	416	3	US-08-946-329A-17	Sequence 17, Appl
40	41	31.8	416	3	US-08-567-357A-17	Sequence 17, Appl
41	41	31.8	416	3	US-08-729-743A-17	Sequence 17, Appl
42	41	31.8	416	3	US-08-349-498-17	Sequence 25, Appl
43	41	31.8	416	4	US-09-325-256-25	Sequence 17, Appl
44	41	31.8	416	5	PCT-US95-15463-17	Sequence 17, Appl
45	41	31.8	416	5	PCT-US95-15923-17	Sequence 17, Appl

## ALIGNMENTS

```

RESULT 1
US-08-475-886-2
; Sequence 2, Application US/08475886A
; Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
EARLIER FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232
EARLIER FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-08-475-886-2
Query Match 100.0%; Score 129; DB 3; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2.9e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WLNPKKINLADRMGLSGVOEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVOEIKEQ 980
RESULT 2
US-08-475-886-4
; Sequence 4, Application US/08475886A
; Patent No. 6113912
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PURCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/08/475,886A
CURRENT FILING DATE: 1995-06-07
EARLIER APPLICATION NUMBER: 07/947,338
EARLIER FILING DATE: 1992-09-18
EARLIER APPLICATION NUMBER: 08/397,232

```

EARLIER FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
US-08-475-886-4

Query Match 100.0%; Score 129; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2,9e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 3  
US-08-475-886-6  
Sequence 6, Application US/08475886A  
Patent No. 6113912  
GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 202642620S2  
CURRENT APPLICATION NUMBER: US/08/475,886A  
EARLIER FILING DATE: 1995-06-07  
EARLIER APPLICATION NUMBER: 07/947,338  
EARLIER FILING DATE: 1992-09-18  
EARLIER APPLICATION NUMBER: 08/397,232  
EARLIER FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-08-475-886-6

Query Match 100.0%; Score 129; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2,9e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 4  
US-08-397-232-2  
Sequence 2, Application US/08397232A  
Patent No. 6180110  
GENERAL INFORMATION:  
APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 202642620S1  
CURRENT APPLICATION NUMBER: US/08/397,232A  
EARLIER FILING DATE: 1995-04-17  
EARLIER APPLICATION NUMBER: 07/947,338  
EARLIER FILING DATE: 1992-09-18  
EARLIER APPLICATION NUMBER: PCT/US93/08610  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 2227

TYPE: PRT  
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-08-397-232-2

Query Match 100.0%; Score 129; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2,9e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 5  
US-08-397-232-4  
Sequence 4, Application US/08397232A  
Patent No. 6180110  
GENERAL INFORMATION:

APPLICANT: FUNKHOUSER, ANN W  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 202642620S1  
CURRENT APPLICATION NUMBER: US/08/397,232A  
EARLIER FILING DATE: 1995-04-17  
EARLIER APPLICATION NUMBER: 07/947,338  
EARLIER FILING DATE: 1992-09-18  
EARLIER APPLICATION NUMBER: PCT/US93/08610  
EARLIER FILING DATE: 1993-09-17  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-08-397-232-4

Query Match 100.0%; Score 129; DB 3; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 2,9e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 6  
US-09-171-387-2  
Sequence 2, Application US/09171387  
Patent No. 6280734  
GENERAL INFORMATION:

APPLICANT: RAYCHAUDHURI, GOPA;  
EMERSON, SUZANNE, U.;  
PURCELL, ROBERT, H.  
TITLE OF INVENTION: SIMIAN-HUMAN HAV  
HAVING A CHIMERIC 2C PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.  
STREET: 345 PARK AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10154  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: MICROSOFT WORD 97  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/171,387  
FILING DATE: 24-Mar-1999

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/06506
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US60/015,642
FILING DATE: 19-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: William S. Feller
REGISTRATION NUMBER: 26,728
REFERENCE/DOCKET NUMBER: 2026-42290S1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
INFORMATION FOR SEQ ID NO: 2
SEQUENCE CHARACTERISTICS:
LENGTH: 2227 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
09-171-387-2

```

```

Query Match          100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2,9e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

```

```

RESULT 7
US-09-653-499-2
Sequence 2, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 2227
TYPE: PRT
ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-09-653-499-2

```

```

Query Match          100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2,9e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

```

```

RESULT 8
US-09-653-499-4
Sequence 4, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U

```

```

APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated HAV (Pass 35), strain HM-175
US-09-653-499-4

```

```

Query Match          100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2,9e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

```

```

RESULT 9
US-09-653-499-6
Sequence 6, Application US/09653499
Patent No. 6423318
GENERAL INFORMATION:
APPLICANT: FUNKHOUSER, ANN W
APPLICANT: EMERSON, SUZANNE U
APPLICANT: PORCELL, ROBERT H
APPLICANT: D'HONDT, ERIC
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
FILE REFERENCE: 20264262US2
CURRENT APPLICATION NUMBER: US/09/653,499
CURRENT FILING DATE: 2000-08-31
PRIOR APPLICATION NUMBER: 08/475,886
PRIOR FILING DATE: 1999-08-09
PRIOR APPLICATION NUMBER: 07/947,338
PRIOR FILING DATE: 1992-09-18
PRIOR APPLICATION NUMBER: 08/397,232
PRIOR FILING DATE: 1995-03-10
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 2227
TYPE: PRT
ORGANISM: Attenuated (4380) HAV, strain HM-175
US-09-653-499-6

```

```

Query Match          100.0%; Score 129; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 2,9e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 WLNPKKINLADRMGLSGVQEIKEQ 25
Db 956 WLNPKKINLADRMGLSGVQEIKEQ 980

```

```

RESULT 10
US-09-248-528-17
Sequence 17, Application US/09248528C
Patent No. 6153415
GENERAL INFORMATION:
APPLICANT: Ortel, Patrick J
APPLICANT: Padmakumar, Rugmini
APPLICANT: Kim, Sang H

```

```
;; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
;; FILE REFERENCE: MSU 4.1-401
;; CURRENT APPLICATION NUMBER: US/09/248,528C
;; EARLIER FILING DATE: 1999-02-10
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: Patentin Ver. 2.0
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Bacillus smithii
US-09-248-528-17

Query Match      38.0%; Score 49; DB 3; Length 229;
Best Local Similarity 41.7%; Pred. No. 1.4;
Matches 10; Conservative 6; Mismatches 6; Indels 2; Gaps 1;
```

```
QY      1 WLNPKKINLADRL--GLSGVOEI 22
      111 : : : 111 : : :
      113 WENPKLVKVEKALLEGLSPVREV 136
```

```
RESULT 11
US-09-549-108-17
;; Sequence 17, Application US/09549108
;; Patent No. 6214603
;; GENERAL INFORMATION:
;; APPLICANT: OriTel, Patrick J
;; APPLICANT: Padmakumar, Rugmini
;; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
;; FILE REFERENCE: MSU 4.1-486
;; CURRENT APPLICATION NUMBER: US/09/549,108
;; PRIOR FILING DATE: 2000-04-13
;; PRIOR APPLICATION NUMBER: 60/083,485
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 09/248,528
;; PRIOR FILING DATE: 1999-02-10
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 17
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Bacillus smithii
US-09-549-108-17

Query Match      38.0%; Score 49; DB 3; Length 229;
Best Local Similarity 41.7%; Pred. No. 1.4;
Matches 10; Conservative 6; Mismatches 6; Indels 2; Gaps 1;
```

```
QY      1 WLNPKKINLADRL--GLSGVOEI 22
      111 : : : 111 : : :
      113 WENPKLVKVEKALLEGLSPVREV 136
```

```
RESULT 12
US-09-549-111-17
;; Sequence 17, Application US/09549111
;; Patent No. 6228633
;; GENERAL INFORMATION:
;; APPLICANT: OriTel, Patrick J
;; APPLICANT: Padmakumar, Rugmini
;; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
;; FILE REFERENCE: MSU 4.1-489
;; CURRENT APPLICATION NUMBER: US/09/549,111
;; PRIOR FILING DATE: 2000-04-13
;; PRIOR APPLICATION NUMBER: 60/083,485
;; PRIOR FILING DATE: 1998-04-29
```

```
;; PRIOR APPLICATION NUMBER: 09/248,528
;; PRIOR FILING DATE: 1999-02-10
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 17
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Bacillus smithii
US-09-549-111-17
```

```
Query Match      38.0%; Score 49; DB 3; Length 229;
Best Local Similarity 41.7%; Pred. No. 1.4;
Matches 10; Conservative 6; Mismatches 6; Indels 2; Gaps 1;
```

```
QY      1 WLNPKKINLADRL--GLSGVOEI 22
      111 : : : 111 : : :
      113 WENPKLVKVEKALLEGLSPVREV 136
```

```
RESULT 13
US-09-549-106-17
;; Sequence 17, Application US/09549106
;; Patent No. 6242242
;; GENERAL INFORMATION:
;; APPLICANT: OriTel, Patrick J
;; APPLICANT: Padmakumar, Rugmini
;; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
;; FILE REFERENCE: MSU 4.1-487
;; CURRENT APPLICATION NUMBER: US/09/549,106
;; PRIOR FILING DATE: 2000-04-13
;; PRIOR APPLICATION NUMBER: 60/083,485
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 09/248,528
;; PRIOR FILING DATE: 1999-02-10
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 17
;; LENGTH: 229
;; TYPE: PRT
;; ORGANISM: Bacillus smithii
US-09-549-106-17

Query Match      38.0%; Score 49; DB 3; Length 229;
Best Local Similarity 41.7%; Pred. No. 1.4;
Matches 10; Conservative 6; Mismatches 6; Indels 2; Gaps 1;
```

```
QY      1 WLNPKKINLADRL--GLSGVOEI 22
      111 : : : 111 : : :
      113 WENPKLVKVEKALLEGLSPVREV 136
```

```
RESULT 14
US-09-550-394-17
;; Sequence 17, Application US/09550394
;; Patent No. 6287828
;; GENERAL INFORMATION:
;; APPLICANT: OriTel, Patrick J
;; APPLICANT: Padmakumar, Rugmini
;; TITLE OF INVENTION: Method for Producing Amide Compounds Using a Nitrile
;; FILE REFERENCE: MSU 4.1-488
;; CURRENT APPLICATION NUMBER: US/09/550,394
;; PRIOR FILING DATE: 2000-04-14
;; PRIOR APPLICATION NUMBER: 60/083,485
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 09/248,528
;; PRIOR FILING DATE: 1999-02-10
;; NUMBER OF SEQ ID NOS: 18
;; SOFTWARE: Patentin Ver. 2.0
;; SEQ ID NO 17
```

LENGTH: 229  
 TYPE: PRT  
 ORGANISM: Bacillus smith1  
 US-09-550-394-17

Query Match 38.0%; Score 49; DB 3; Length 229;  
 Best Local Similarity 41.7%; Pred. No. 1.4;  
 Matches 10; Conservative 6; Mismatches 6; Indels 2; Gaps 1;

OY 1 WLPKKNLADRL--GLSGVOEI 22  
 | | | : : : | | | : | :  
 Db 113 WENPKLVKVEKALLEGISPVREV 136

RESULT 15  
 US-09-252-991A-17077  
 Sequence 17077, Application US/09252991A  
 Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

NUMBER OF SEQ ID NOS: 33142

SEQ ID NO 17077

LENGTH: 699

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-17077

Query Match 36.8%; Score 47.5; DB 4; Length 699;  
 Best Local Similarity 45.5%; Pred. No. 1.0;  
 Matches 10; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

OY 2 LNPKNLADRLGLSGVOEI 23  
 | | | | : : : | | : | :  
 Db 310 LNPKNL-IDMVIIGLAGNDAR 330

Search completed: October 1, 2003, 10:06:35  
 Job time : 16.8824 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 1, 2003, 10:02:52 ; Search time 30.1471 seconds  
(without alignments)  
131.201 Million cell updates/sec

Title: US-09-171-432a-48  
Perfect score: 129  
Sequence: 1 WLNPKINLADRLGLSGVQEIKEQ 25

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 587654 seqs, 158212981 residues  
Total number of hits satisfying chosen parameters: 587654

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PC1\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PC1S\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09C\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	129	100.0	352	15 US-10-272-459-45	Sequence 45, Appl
2	129	100.0	980	15 US-10-272-459-41	Sequence 41, Appl
3	129	100.0	2227	15 US-09-929-955-12	Sequence 12, Appl
4	129	100.0	2227	14 US-10-104-966-12	Sequence 12, Appl
5	129	100.0	2227	14 US-10-135-988-2	Sequence 2, Appl
6	129	100.0	2227	14 US-10-135-988-4	Sequence 4, Appl
7	129	100.0	2227	14 US-10-135-988-6	Sequence 6, Appl
8	48.5	37.6	1987	15 US-10-133-382-6	Sequence 6, Appl
9	48.5	37.6	2013	15 US-10-133-382-2	Sequence 2, Appl
10	48.5	37.6	2014	15 US-10-133-382-8	Sequence 8, Appl
11	48.5	37.6	2040	11 US-10-133-382-4	Sequence 4, Appl
12	45	34.9	309	11 US-09-510-332-67	Sequence 67, Appl
13	45	34.9	325	10 US-09-764-864-1535	Sequence 1535, Ap
14	45	34.9	357	9 US-09-938-803-16	Sequence 16, Appl
15	45	34.9	371	10 US-09-764-864-1118	Sequence 1118, Ap

16	45	34.9	622	15 US-10-156-761-13093	Sequence 13093, A
17	44	34.1	1770	10 US-09-801-368-298	Sequence 298, App
18	43	33.3	289	15 US-10-286-264-110	Sequence 110, App
19	43	33.3	358	9 US-09-861-451A-50	Sequence 50, Appl
20	43	33.3	426	15 US-10-156-761-9490	Sequence 9490, Ap
21	43	33.3	495	9 US-09-815-242-11215	Sequence 11215, A
22	42	32.6	166	9 US-09-815-242-4942	Sequence 4942, Ap
23	42	32.6	169	9 US-09-815-242-10531	Sequence 10531, A
24	42	32.6	185	10 US-09-250-611-11	Sequence 11, Appl
25	42	32.6	299	9 US-09-825-882-16	Sequence 16, Appl
26	42	32.6	299	12 US-10-017-161-1832	Sequence 1832, Ap
27	42	32.6	405	14 US-10-029-180-78	Sequence 78, Appl
28	42	32.6	434	14 US-10-029-180-112	Sequence 112, App
29	42	32.6	608	15 US-10-153-668-196	Sequence 196, Appl
30	42	32.6	638	15 US-10-216-209-95	Sequence 95, Appl
31	42	32.6	1250	10 US-09-801-368-164	Sequence 364, App
32	42	32.6	1404	9 US-09-811-045A-1	Sequence 1, Appl
33	42	32.6	2121	10 US-09-995-542-3	Sequence 3, Appl
34	42	32.6	2167	10 US-09-995-542-2	Sequence 2, Appl
35	42	32.6	5701	9 US-09-864-761-37319	Sequence 37319, A
36	41	31.8	184	10 US-09-250-611-13	Sequence 13, Appl
37	41	31.8	184	15 US-10-205-823-417	Sequence 417, App
38	41	31.8	184	15 US-10-177-293-455	Sequence 455, App
39	41	31.8	197	15 US-10-146-473-74	Sequence 74, Appl
40	41	31.8	299	12 US-10-298-638-22	Sequence 22, Appl
41	41	31.8	362	10 US-09-764-864-896	Sequence 896, App
42	41	31.8	366	10 US-09-870-759-72	Sequence 72, Appl
43	41	31.8	366	12 US-09-751-708A-72	Sequence 72, Appl
44	41	31.8	366	12 US-10-281-478-10	Sequence 10, Appl
45	41	31.8	414	12 US-10-032-585-7663	Sequence 7663, Ap

## ALIGNMENTS

RESULT 1  
US-10-272-459-45  
; Sequence 45, Application US/10272459  
; Publication No. US20030124517A1  
; GENERAL INFORMATION:  
; APPLICANT: PICQUANTES, Sergio  
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT  
; FILE REFERENCE: PP1955.002 / 2301-17955  
; CURRENT APPLICATION NUMBER: US/10/272,459  
; NUMBER OF SEQ ID NOS: 50  
; SOFTWARE: Patentlin Ver. 2.0  
; SEQ ID NO 45  
; LENGTH: 352  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein  
; OTHER INFORMATION: of 38.8 kDa  
US-10-272-459-45

Query Match 100.0%; Score 129; DB 15; Length 352;  
Best Local Similarity 100.0%; Pred. No. 5.4e-13;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKINLADRLGLSGVQEIKEQ 25  
|||||  
Db 328 WLNPKINLADRLGLSGVQEIKEQ 352

RESULT 2  
US-10-272-459-41  
; Sequence 41, Application US/10272459  
; Publication No. US20030124517A1  
; GENERAL INFORMATION:  
; APPLICANT: PICQUANTES, Sergio  
; TITLE OF INVENTION: HEPATITIS A VIRUS NUCLEOTIDE SEQUENCES, RECOMBINANT

```

; TITLE OF INVENTION: PROTEINS AND USES THEREOF
; FILE REFERENCE: PPI7955.002 / 2301-17955
; CURRENT APPLICATION NUMBER: US/10/272,459
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 980
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: recombinant protein
US-10-272-459-41

Query Match          100.0%; Score 129; DB 15; Length 980;
Best Local Similarity 100.0%; Pred. No. 1.8e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25
   |||||||||||||||||||||
Db 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 3
US-09-929-955-12
; Sequence 12, Application US/09929955
; Patent No. US20020136740A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: METHODS OF USE THEROF
; TITLE OF INVENTION: METHODS OF USE THEROF
; FILE REFERENCE: TRIPEP.23AUS2
; CURRENT APPLICATION NUMBER: US/09/929,955
; CURRENT FILING DATE: 2001-08-15
; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/225,767
; PRIOR FILING DATE: 2000-08-17
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
09-929-955-12

Query Match          100.0%; Score 129; DB 10; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.6e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25
   |||||||||||||||||||||
Db 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 4
US-10-104-966-12
; Sequence 12, Application US/10104966
; Publication No. US20020155124A1
; GENERAL INFORMATION:
; APPLICANT: Matti Salberg
; APPLICANT: Catharina Hultgren
; TITLE OF INVENTION: VACCINES CONTAINING RIBAVIRIN AND
; FILE REFERENCE: TRIPEP.23AUSC1
; CURRENT APPLICATION NUMBER: US/10/104,966
; CURRENT FILING DATE: 2002-03-22
```

```

; PRIOR APPLICATION NUMBER: 09/705,547
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: 60/229,175
; PRIOR FILING DATE: 2000-08-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Hepatitis A virus sequence
US-10-104-966-12

Query Match          100.0%; Score 129; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.6e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25
   |||||||||||||||||||||
Db 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 5
US-10-135-988-2
; Sequence 2, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PORCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
; PRIOR FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2227
; TYPE: PRT
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175
US-10-135-988-2

Query Match          100.0%; Score 129; DB 14; Length 2227;
Best Local Similarity 100.0%; Pred. No. 4.6e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WLNPKKINLADRMGLSGVOEIKQ 25
   |||||||||||||||||||||
Db 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 6
US-10-135-988-4
; Sequence 4, Application US/10135988
; Publication No. US20020176869A1
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PORCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US3
; CURRENT APPLICATION NUMBER: US/10/135,988
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 07/947,338
; PRIOR FILING DATE: 1992-09-18
; PRIOR APPLICATION NUMBER: 08/397,232
```



PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated HAV (Pass 35), strain HM-175  
US-10-135-988-4

Query Match 100.0%; Score 129; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 4, 6e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WLNPKKINLADRMGLSGVOEIKQ 25  
DB 956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 7  
US-10-135-988-6  
Sequence 6, Application US/10135988  
Publication No. US20020176869A1  
GENERAL INFORMATION:  
APPLICANT: PUNKHouser, ANN M  
APPLICANT: EMERSON, SUZANNE U  
APPLICANT: PURCELL, ROBERT H  
APPLICANT: D'HONDT, ERIC  
TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
FILE REFERENCE: 20264262U3  
CURRENT APPLICATION NUMBER: US/10/135, 988  
CURRENT FILING DATE: 2002-04-29  
PRIOR APPLICATION NUMBER: 07/947,338  
PRIOR FILING DATE: 1992-09-18  
PRIOR APPLICATION NUMBER: 08/397,232  
PRIOR FILING DATE: 1995-03-10  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 2227  
TYPE: PRT  
ORGANISM: Attenuated (4380) HAV, strain HM-175  
US-10-135-988-6

Query Match 100.0%; Score 129; DB 14; Length 2227;  
Best Local Similarity 100.0%; Pred. No. 4, 6e-12;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 WLNPKKINLADRMGLSGVOEIKQ 25  
956 WLNPKKINLADRMGLSGVOEIKQ 980

RESULT 8  
US-10-132-382-6  
Sequence 6, Application US/10132382  
Publication No. US20030045699A1  
GENERAL INFORMATION:  
APPLICANT: WEISS, BERTRAM  
TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS  
FILE REFERENCE: SCH-1811  
CURRENT APPLICATION NUMBER: US/10/132,382  
CURRENT FILING DATE: 2002-04-26  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 6  
LENGTH: 1987  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-132-382-6

Query Match 37.6%; Score 48.5; DB 15; Length 1987;  
Best Local Similarity 50.0%; Pred. No. 79;  
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

OY 2 LNPCKINLADRMGLSGVOEIKQ 24  
DB 275 INLRKLNLSDNHGLGELPGVOSSDE 298

RESULT 9  
US-10-132-382-2  
Sequence 2, Application US/10132382  
Publication No. US20030045699A1  
GENERAL INFORMATION:  
APPLICANT: WEISS, BERTRAM  
TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS  
FILE REFERENCE: SCH-1811  
CURRENT APPLICATION NUMBER: US/10/132,382  
CURRENT FILING DATE: 2002-04-26  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 2013  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-132-382-2

Query Match 37.6%; Score 48.5; DB 15; Length 2013;  
Best Local Similarity 50.0%; Pred. No. 80;  
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

OY 2 LNPCKINLADRMGLSGVOEIKQ 24  
DB 301 INLRKLNLSDNHGLGELPGVOSSDE 324

RESULT 10  
US-10-132-382-8  
Sequence 8, Application US/10132382  
Publication No. US20030045699A1  
GENERAL INFORMATION:  
APPLICANT: WEISS, BERTRAM  
TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS  
FILE REFERENCE: SCH-1811  
CURRENT APPLICATION NUMBER: US/10/132,382  
CURRENT FILING DATE: 2002-04-26  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 8  
LENGTH: 2014  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-132-382-8

Query Match 37.6%; Score 48.5; DB 15; Length 2014;  
Best Local Similarity 50.0%; Pred. No. 80;  
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

OY 2 LNPCKINLADRMGLSGVOEIKQ 24  
DB 302 INLRKLNLSDNHGLGELPGVOSSDE 325

RESULT 11  
US-10-132-382-4  
Sequence 4, Application US/10132382  
Publication No. US20030045699A1  
GENERAL INFORMATION:  
APPLICANT: WEISS, BERTRAM  
TITLE OF INVENTION: MEMBRANE RECEPTORS FOR STEROIDS AND STEROLS  
FILE REFERENCE: SCH-1811  
CURRENT APPLICATION NUMBER: US/10/132,382  
CURRENT FILING DATE: 2002-04-26  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4

LENGTH: 2040  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-132-382-4

Query Match 37.6%; Score 48.5; DB 15; Length 2040;  
Best Local Similarity 50.0%; Pred. No. 81;  
Matches 12; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

2 LNPKKINLADRMGLSGVOEIKE 24  
Db 328 LNPKKINLSDNHLGELPEGVSSDE 351

RESULT 12  
US-09-510-332-67  
Sequence 67, Application US/09510332  
Publication No. US20030022278A1  
GENERAL INFORMATION:  
APPLICANT: Zuker, Charles S.  
APPLICANT: Adler, Jon Elliott  
APPLICANT: Ryba, Nick  
APPLICANT: Mueller, Ken  
APPLICANT: Hoon, Mark  
APPLICANT: The Regents of the University of California  
TITLE OF INVENTION: T2R, a No. US20030022278A1 Family of Taste Receptors  
FILE REFERENCE: 02307E-098010US  
CURRENT APPLICATION NUMBER: US/09/510,332  
CURRENT FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 09/393,634  
PRIOR FILING DATE: 1999-09-10  
NUMBER OF SEQ ID NOS: 172  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 67  
LENGTH: 309  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: human T2R41 (hGR41)  
NAME/KEY: MOD\_RES  
LOCATION: (253)  
OTHER INFORMATION: Xaa - any amino acid  
US-09-510-332-67

Query Match 34.9%; Score 45; DB 11; Length 309;  
Best Local Similarity 36.4%; Pred. No. 34;  
Matches 8; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

1 WLNPKKINLADRMGLSGVOEIKE 22  
35 WLNTRKISSADQIILTAIAVSRY 56

RESULT 13  
US-09-764-864-1535  
Sequence 1535, Application US/09764864  
Patent No. US20020132753A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT23  
CURRENT APPLICATION NUMBER: US/09/764,864  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 1792  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1535  
LENGTH: 325  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (309)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (322)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-864-1535

Query Match 34.9%; Score 45; DB 10; Length 325;  
Best Local Similarity 45.5%; Pred. No. 37;  
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

3 NPKKINLADRMGLSGVOEIKE 24  
Db 210 NPKKILHTIGMLVLLSEEEIQO 231

RESULT 14  
US-09-938-803-16  
Sequence 16, Application US/09938803  
Patent No. US20020076762A1  
GENERAL INFORMATION:  
APPLICANT: Yue, Henry  
APPLICANT: Tang, Y. Tom  
APPLICANT: Lal, Preeti  
APPLICANT: Reddy, Roopa  
APPLICANT: Baughn, Mariah R.  
APPLICANT: Yang, Junning  
APPLICANT: Azimzai, Yalda  
TITLE OF INVENTION: FULL-LENGTH EXPRESSED GENETIC MARKERS  
FILE REFERENCE: PF-0695 US  
CURRENT APPLICATION NUMBER: US/09/938,803  
CURRENT FILING DATE: 2000-08-24  
PRIOR APPLICATION NUMBER: 09/311,894  
PRIOR FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 40  
SOFTWARE: PERL Program  
SEQ ID NO 16  
LENGTH: 357  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc.feature  
OTHER INFORMATION: Incyte Clone 2805526  
US-09-938-803-16

Query Match 34.9%; Score 45; DB 9; Length 357;  
Best Local Similarity 45.5%; Pred. No. 41;  
Matches 10; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

3 NPKKINLADRMGLSGVOEIKE 24  
Db 174 NPKKILHTIGMLVLLSEEEIQO 195

RESULT 15  
US-09-764-864-1118  
Sequence 1118, Application US/09764864  
Patent No. US20020132753A1  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PT23  
CURRENT APPLICATION NUMBER: US/09/764,864  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 1792  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1118  
LENGTH: 371  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-764-864-1118

Query Match 34.9%; Score 45; DB 10; Length 371;

Best Local Similarity 45.5%; Pred. No. 43;  
Matches 10; Conservative 5; Mismatches

7; Indels 0; Gaps 0;

OY 3 NPKINLADRMGLSGVOEIKE 24  
||||:| ||| :||:  
Db 188 NPKILHLTIGMLVLLSEETIOO 209

Search completed: October 1, 2003, 10:37:54  
Job time : 30.1471 secs

**THIS PAGE BLANK (USPTO)**